#Fragilemasculinity: The Roles of Masculinity Threat and Anonymity in Men's Endorsement and Perpetration of Online Gender Harassment

by

Jennifer D. Rubin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology and Women's Studies)
in the University of Michigan
2019

Doctoral Committee:

Associate Professor Terri D. Conley, Chair
Associate Professor Sonya Dal Cin
Professor Nicole Ellison
Associate Professor Sara I. McClelland

Jennifer Rubin

jdrrubin@umich.edu

ORCID iD: 0000-0002-6649-6606

DEDICATION

To Dahlia Josephine
I've loved you since before you were born.
To Matthew
My unwavering support and love.
To Debbie and Don Rubin
Without whom none of my success would be possible.

ACKNOWLEDGMENTS

I have received encouragement, mentorship, and support from a community of friends and scholars. Thank you to the IDEAL family for providing me guidance during my undergraduate years at UC Berkeley. The support I received from Cedric Brown, Genevieve Negrón-Gonzales, and Mitch and Frida Kapor helped me realize my dream of pursing a graduate degree. I would also like to thank Drs. Mel Chen and Barbara Barnes for nurturing my commitment to Gender and Women's Studies as critical to undergraduate education. A big thank you to Dr. David Frost for his mentorship and guidance during my time at San Francisco State University. My dissertation committee of Drs. Terri Conley, Sara McClelland, Sonya Dal Cin, and Nicole Ellison have provided me with invaluable suggestions that strengthened this research and future projects. And a very special thank you to Dr. Terri Conley for being my springboard for research ideas and my unwavering support. I am so happy to have you as a friend and mentor.

I would like to acknowledge the friendships I have built while at San Francisco State and University of Michigan. Thank you to the following people: Sara Chadwick, Jes Matsick, Logan Casey, Danielle Shapiro, Jesse Chandler, Rita Seabrook, Aayat Ali, Jad Raad, Ali Ziegler, Verónica Rabelo, Katy Holland, and Cat Forrester. My outstanding research assistants have made this dissertation possible: Rachel Cultice, Mary Kruk, Shoshana Kaplan, Rizzo Nichols, Morgan Perry, Charles Vazquez, and Katherine Robbins. I am fully indebted to the Personality and Social Contexts Psychology area, Gender and Feminist Psychology area, and the Women's Studies Department for always listening and giving me words of encouragement.

I would also like to express my sincere gratitude to Dr. Joanne Bailey, teacher and mentor extraordinaire. Your passionate approach to teaching Perspectives in Women's Health is contagious. And thank you for opening your home in Fall 2017 so I could complete data collection for my dissertation. I look back fondly on our conversations after work and sharing a meal together.

Thank you to my partner, Matthew Snyder. You moved across the country, leaving our rent-controlled apartment in San Francisco, so I could pursue a PhD at Michigan. You have been my biggest support, and I love you dearly. Six years later, we are almost done with graduate school. What a journey it has been!

Finally, thank you to Debbie and Don Rubin, without whom none of my success would be possible. Although both of you never had the opportunity to attend college, you taught me that education was key to success. One of Don's final wishes before his death was that I pursue higher education— and look how far I've come, Dad. I share this tremendous accomplishment with both of you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

The present studies examined whether masculinity threat and anonymity impacted men's support and use of gender harassment in social media. Using survey and experimental methods, this research evaluated men's endorsement and perpetration of gender harassment on Twitter and Facebook across three studies (N_{Study1} = 258; N_{Study2} = 94; N_{Study3}=216). Study 1 found that men's stress about masculinity failures was negatively related to harassment endorsement on Twitter. In Study 2, public shame about masculinity failures played a key role in men's *responses* to masculinity threat. Following masculinity threat, men who expressed concern that others would perceive them as insufficiently masculine reported greater harassment endorsement on Twitter. Study 3 showed that the relationship between masculinity threat and anonymity impacted men's *behaviors* on Facebook. Masculinity-threatened men who were anonymous shared more sexist memes compared to masculinity-threatened men who were identified. Additionally, among masculinity-threatened men, men who reported fearing backlash about their masculinity failures shared more sexist memes. Taken together, these findings demonstrate that gender and anonymity are important factors in men's endorsement and perpetration of gender harassment in social media.

CHAPTER 1

Introduction

Women navigate an unprecedented amount of harassment online. The past few years have produced several noteworthy cases in social media, such as #Gamergaters' sexist attacks on women in gaming communities and Reddit forums that promote misogyny (Mantilla, 2015; Poland, 2016). Across these high-profile incidences, harassment took the form of negative comments on blogs and social networking sites (Franks, 2012; Jane, 2014a); cyberstalking (Barak, 2005); and violent threats that often targeted users on the basis of their gender (Henry & Powell, 2015; Jane, 2016). In many respects, online harassment is a proxy for a cultural war over inclusion of women— and as such, it is a battle for who has power to speak online. This ultimately gives rise to new questions: Do women experience distinct forms of online harassment? What motivates male perpetrators to engage in online harassment targeted at women?

Online harassment refers to a range of abusive behaviors facilitated by technology platforms. These behaviors include but are not limited to flaming (the use of inflammatory language, name calling, or insults); doxing (the public release of identifiable information, such as a home address, social security number or phone number); and public shaming (the use of social media sites to humiliate a target or damage their reputation, Blackwell, Dimond, Schoenebeck, & Lampe, 2017). Although women and men experience similar rates of harassment online, women are disproportionately affected by more extreme violations, including being sexually harassed

and physically threatened (Duggan, 2014). *Gender harassment* occurs in these venues when an individual is belittled or demeaned based on her gender (Leskinen, Rabelo, & Cortina, 2015; Woodzicka & LaFrance, 2001). These violations may include the transmission of gender-degrading materials, sexist jokes and slurs, and offensive comments that evoke gender stereotypes (Leskinen & Cortina, 2014). Online gender harassment has detrimental effects—women frequently report emotional and physical distress, self-censoring online, and increased safety concerns following incidences of mistreatment (Poland, 2016).

The decision to harass women in social media is seldom unplanned. Male perpetrators employ the same stereotypes and derailing tactics that appear offline to drive women away, especially from positions of authority and power (Jane, 2014b; Megarry, 2014). The breadth of this aggression has received attention not only from psychologists, but in popular press, where it has been widely critiqued as a consequence of trolling culture (Chu, 2014; Hess, 2014). Largely absent from these conversations is the role of gender, and in particular *threats to masculinity,* in relation to online harassment. Some men, "especially men grappling with real threats to their social well-being on other fronts," may perceive the progress of women as "spelling their own masculine doom" (Faludi, 2009, pp. 10–11). Men's motivations to gender harass may therefore be driven by a desire to "put women, because they are women, back in their rightful place" (Chemaly, 2014).

Scholars have noted that the incursion of women into what were exclusively male spaces may inform efforts of harassers online (Lumsden & Morgan, 2017; Salter, 2017). In particular, some men perceive the heightened visibility of women's voices in social media—such as articulations of popular feminism on Twitter and Facebook— as challenges to the gender status quo and respond with silencing tactics (Banet-Weiser & Miltner, 2016). Online environments are

particularly suited for facilitating gender harassment because online communication tends to loosen normative constraints of behavior that are otherwise unacceptable "offline" (Lea & Spears, 1991; Suler, 2004). Given that public action is important in restoring threatened masculinity (Vandello & Bosson, 2013), social media can provide men with ample opportunities to display demonstrations of proof.

In this research, I examine the relationship between masculinity and sociotechnical affordances in men's endorsement and perpetration of gender harassment. Using survey and experimental methods, I propose that men's use of online gender harassment may effectively appease threats to masculinity. Anonymity afforded by online communication can exacerbate these effects, since the impression of being anonymous makes it easier to engage in harassment (Fox, Cruz & Lee, 2015). Technology therefore enables harassment to thrive, yet men's endorsement and perpetration of harassment are grounded in "offline" gender norms and masculinity expectations.

The present research extends psychological research on men and masculinity(s) in two ways. First, I examine how masculinity threat may drive men to endorse online gender harassment, with an understanding that some men are particularly sensitive to threat given individual differences. Whereas most masculinity threat research does not include moderators, this research considers how men's perceptions of their masculinity impact attitudes and behaviors. Second, I explore the relationship between sociotechnical affordances— structural features of computer technology that enable or constrain certain actions— and behavioral outcomes. In particular, I investigate the relationship between anonymity and likelihood to gender harass following threats to masculinity.

This research also foregrounds a feminist analysis in psychological approaches to online harassment. Some researchers have explored the sexist nature of online abuse, yet many have not acknowledged—let alone prioritized—gender in their analysis (Megarry, 2014). Online harassment is often seen as a pattern of mistreatment that *all* Internet users encounter (Citron, 2014). Rather than frame these behaviors as an idiosyncratic practice that emerged with the Internet, I situate online environments as contexts that reflect "offline" norms surrounding masculinity. I focus on one group of women in particular: women who advocate for gender equality on Twitter and Facebook. In doing this, I aim to develop insight into consequences for women who challenge the gender status quo online.

To fully consider these issues, I join psychological approaches to masculinity with scholarship from women's studies and communication studies. I begin with an overview of masculinity threat research, individual differences in masculinity norms, and intersections between race and masculinity. Next, I explore why women who are in positions of authority contend with amplified mistreatment online, especially when they intervene in male-dominated arenas or advocate for gender equality. Finally, drawing on literature from communication studies and psychology, I consider the role of anonymity in amplifying online gender harassment.

Psychological Approaches to Masculinities and Manhood

Masculinity describes the practices, behaviors, and expectations culturally associated with (though not limited to) people understood to be male (Coston & Kimmel, 2012; Pascoe & Bridges, 2016). Because these gendered behaviors are important parts of some men's identities, masculinity requires social validation through interaction with others (Vandello & Bosson, 2013; West & Zimmerman, 1987). For example, men learn from an early age the importance of norms

surrounding masculinity and femininity and as a result, they are quick to demonstrate that they understand these roles in social situations (Fenstermaker & West, 2002; Thompson & Bennett, 2015). Masculinity is therefore *relational* because it functions "...as an aspect of a larger [gender] structure" and as a result, has no meaning outside its relationship to femininity (Connell & Connell, 2005, p. 67).

Masculinity does not constitute a monolithic standard of behavior but rather comprises a range of expectations that Connell (1985) has termed *hegemonic masculinity*. Hegemonic masculinity refers to current social structures that legitimate men's dominant social position over women, and other gender identities, which are understood as "feminine" in a given culture (Connell & Connell, 2005). Though hegemonic masculinity varies with geographic and historical context (Thompson, Pleck, & Ferrera, 1992), most definitions situate aggression, risk-taking, competitiveness, assertiveness, and toughness as central to masculinity (Kimmel, 2008; Levant, Rankin, Williams, Hasan, & Smalley, 2010; Pascoe, 2011). Importantly, hegemonic masculinity works to legitimate gender inequality by reconstituting hierarchical relationships between (and among) women and men (Connell & Messerschmidt, 2005; Donaldson, 1993). Because femininity is less valued than masculinity, the stakes associated with maintaining hegemonic masculinity are therefore high for men since it validates their position of power within the existing social hierarchy (Dahl, Vescio, & Weaver, 2015).

Regardless of contextually-specific definitions of masculinity, cultures around the world perceive manhood as a social status that must be earned and can be lost (Gilmore, 1990). Vandello and Bosson (2013) contended that although men hold greater structural power than women in most cultures, manhood (relative to womanhood) is a precarious status because a narrower definition exists for what are acceptable masculine traits. They proposed that much of

men's anxiety about adhering to masculinity norms emerge from three tenets: (1) manhood is widely viewed as a precarious, achieved status that is earned during the social transactions of everyday life; (2) once achieved, manhood can be easily lost or taken away; (3) manhood must be consistently reasserted through public demonstrations, requiring social validation and recognition by others. When men fail to demonstrate the core tenets of manhood, they risk losing their tenuous status and may respond with attempts to restore their masculinity (Berke, Reidy, Miller, & Zeichner, 2017).

Since manhood is an esteemed social status, men may be particularly receptive to situational cues that threaten their masculinity (Moore & Stuart, 2005; Vandello & Bosson, 2013). Because women and men are stereotyped in oppositional terms, masculinity threat may motivate men to act in ways that restore their status, especially when a threat invokes femininity (Brascombe, Ellemers, Spears, & Doosje, 1999). Consistent with this framework, empirical research indicates that a man's masculinity can be threatened when he engages in stereotypically feminine activities such as braiding hair (Bosson, Vandello, Burnaford, Weaver, & Wasti, 2009), when he is told that he possesses feminine attributes (Cohn, Seibert, & Zeichner, 2009) , or when he is outperformed by a woman (Dahl et al., 2015).

A number of studies have shown that men compensate for threatened masculinity through attitudes and behaviors that bolster masculine prototypicality. Following masculinity threat, men are more likely to express antigay prejudice (Glick, Gangl, Gibb, Klumpner, & Weinberg, 2007), blame to female sexual assault survivors (Munsch & Willer, 2012), and endorsement of violent military action (Willer, Rogalin, Conlon, & Wojnowicz, 2013). Some men also engage in more violent behaviors, including demonstrations of physical aggression (Bosson et al., 2009) and harassment (Hunt & Gonsalkorale, 2014). For example, in a series of experimental studies,

Maass and collegues (2003) found that men who had their masculinity threatened engaged in more harassing behavior towards a female target than men who did not have their masculinity threatened. Taken together, these findings suggest that some men act to validate manhood when faced with feelings of inadequacy.

Men may strive to prove or restore manhood to avoid *backlash effects*: social and economic penalties that result from violating gender stereotypes (see Rudman & Fairchild, 2004; Rudman, Moss-Racusin, Phelan, & Nauts, 2012). Men may experience backlash when they violate (1) prescriptions that assign men with upholding traits associated with manhood (e.g., toughness and confidence) and (2) proscriptions of masculinity failures (e.g., weakness and insecurity (Moss-Racusin, Phelan, & Rudman, 2010). Accordingly, men may engage in reparative strategies (e.g., harassment, masculine stereotyped behaviors) to prove their manhood in an effort to avoid backlash from women and other men (Cheryan, Schwartz Cameron, Katagiri, & Monin, 2015). This is especially salient under conditions of threat, as men are motivated to reinstate masculinity.

Individual Differences Perspective. Men's masculinity beliefs may play an important role in their reactions to masculinity threat (Hunt & Gonsalkorale, 2014). Thompson and Bennett (2015) defined masculinity beliefs as expectations about the way men are supposed to behave. For example, some men believe that toughness is central to manhood and behave in ways that display toughness in social situations. Within the context of gender identity threat, research indicates that men who conform to dominant masculinity beliefs may be sensitive to suggestions that they are not traditionally masculine (Weaver & Vescio, 2015). The extent to which men conform to masculine beliefs is therefore of key interest.

Despite progress made by the precarious manhood paradigm in advancing research about masculinity, men's perceptions of their *failures* in masculinity have yet to be integrated into existing threat literature. It is likely that perceptions of oneself as *insufficiently* masculine (i.e., nonconformist to masculine norms) and the *experience of stress* about this discrepancy may also impact behaviors. In other words, what role do men's self-perceived failures in masculinity play in the context of masculinity threat?

The expectation that self-evaluative processes guide men's behaviors falls in line with self-discrepancy theory (Higgins, 1987; Higgins, Klein, & Strauman, 1985). Self-discrepancy theory posits that incongruities arising from the actual self (the representation of attributes that an individual or a perceiver believes an individual possess) and the ought self (the representation of attributes that an individual believes s/he should possess) results in psychological discomfort. Men may experience *gender role discrepancy,* or the perceived failure to live up to masculinity norms, and these self-evaluations may result in stress (Berke et al., 2017). Viewing oneself as insufficiently masculine and distress from this discrepancy may therefore impact attitudes and behaviors in response to a gender threat.

There is reason to suspect that some men at the opposite end of the continuum of masculine norm conformity engage in gender stereotypical behavior. For example, Reidy and colleagues (2016) found that men who endorsed a high level of gender role discrepancy (i.e., being less masculine than average male) and experienced anxiety about this discrepancy reported significantly more sexually risky acts than those who did not experience anxiety. Interestingly, men who endorsed low gender role discrepancy and high discrepancy stress (i.e., highly masculine men) also reported more sexually risky acts. These results indicate that men who

experience stress about being perceived as masculine may be at greater risk for detrimental behaviors, regardless of their high or low conformity to masculine norms.

An aim of the current research was to build on the work of Vandello and Bosson (2013) by examining how men's perceptions of their *masculinity failures* impact endorsement of harassment (Study 1). Conformity to masculinity norms can exacerbate the psychological effects of threat; however, non-conforming men may also experience social pressure to act in stereotypically masculine ways because of their lower social position among other men (Courtenay, 2000). Accordingly, *stress* about being perceived as insufficiently masculine may play an important role in men's behaviors above and beyond men's conformity to masculine norms. This raises the questions: Do men who perceive themselves as fulfilling masculinity ideals behave differently than men who do not perceive themselves as fulfilling masculinity ideals? Do men high in stress about masculinity failures behave differently than men low in stress about masculinity failures?

An additional aim of the current research was to examine the roles of public shame (Study 2) and backlash (Study 3) in men's endorsement and perpetration of harassment. Previous research documents that some men experience concern that others will perceive them as insufficiently masculine when faced with masculinity threat, which I refer to as *public shame* (Rudman & Fairchild, 2004). Men who experience public shame following gender identity threat may be driven to support attitudes that bolster their masculinity compared to men who do not experience public shame. Related, some men who violate gendered expectations report fearing backlash for their masculinity failures. These men may engage in compensatory recovery strategies, such as conforming more to masculinity norms, to avoid backlash (Moss-Racusin & Rudman, 2010). In the current research, I examine how public shame and fearing backlash

moderate the relationship between masculinity threat and outcomes of interest.

Masculinity, Heterosexuality, & Whiteness. The psychology of men and masculinities does not always engage with the interconnectedness of whiteness and masculinity (c.f. Goff, Di Leone, & Kahn, 2012). Experimental research investigating threats to masculinity often situate men— regardless of race and sexuality— as having equal access to hegemonic masculine power (e.g., Cheryan, Schwartz Cameron, Katagiri, & Monin, 2015; Hunt & Gonsalkorale, 2014). Masculinity, however, is not homogeneous; rather, men claim, and sustain, their dominant position in a social hierarchy by demonstrating hegemonic masculinity (Connell, 1985). It is more precise to think of hegemonic masculinity as reflecting a particular position in the social order for privileged men:

> In an important sense there is only one unblushing male in America: a young, married, white, urban, northern, heterosexual, Protestant father of college education, fully employed, of good complexion, weight, and height, and a recent record in sports. Every American male tends to look out upon the world from this perspective ... Any male who fails to qualify in any one of these ways is likely to view himself-during moments at least-as unworthy, incomplete, and inferior. (Goffman, 1963, p. 128)

From a psychological perspective, the intersections between race and sexuality shape behavioral conformity to dominant masculine norms. Kimmel (1994) argued that in the United States, mainstream culture sets "white, middle-class, early middle-aged, heterosexual" as the "standards for other men, against which men are measured and, more often than not, left wanting" (p. 124). Implicit within Kimmel's analysis is how whiteness *and* heterosexual masculinity signify the prototypic man, with marginalized and subordinated men largely invisible. Kimmel (2013) called this the "Goldilocks Dilemma"— masculinities of the other are

perceived as either "too hot" or "too cold," but never "just right." That's not to say that marginalized groups do not perform hegemonic masculinity in order to gain privileges (see Chen, 1999; Cheng, 1999). Rather the hegemonic definition of manhood— that men should behave in ways that demonstrate power, status, and dominance— largely reflect the social privileges available to white, heterosexual men.

Intersections between masculinity and whiteness remain largely absent in the field of psychology and in particular, masculinity threat research. Whiteness is understood to be unmarked and invisible, meaning that it is the standard against which all other differences are measured (Moreton-Robinson, 2000). The invisibility of whiteness, for example, is seen in how psychologists quantitatively measure endorsement and internalization of cultural belief systems about masculinity. These measures largely assess behaviors associated with status, power, and dominance without considering how marginalized men do not share equally in the fruits of these privileges. Given the importance of social location in men's behaviors and attitudes, I limit my investigation to white, heterosexual men within the current research.

Having situated masculinity within a psychological framework, I consider the ways that online contexts reproduce norms about masculinity, femininity, and the relationship between the two. I explore literature from women's studies and psychology to examine (1) how gender manifests in digital environments; and (2) why women who are in positions of authority contend with amplified mistreatment online.

Gender and Online Contexts

Online gender harassment— whether in the form of a Facebook post, Tweet, or Reddit thread—are not isolated from outside events or influences. Rather these expressions reflect beliefs grounded in gendered and racialized dynamics of power aimed at reinforcing male

control of online environments (Fahs & Gohr, 2012). As in offline settings, men are accustomed to dominating online conversations and keeping women's participation low (Poland, 2016). Online harassers use tactics such as derailing (an attempt to disrupt conversation by refocusing the conversations around members of privileged groups) to reinforce sexist attitudes that regard women as invisible (Mantilla, 2015). It is not surprising that some women situate the Internet as having a "boys'-locker-room feel" to it, with male users escalating cases of harassment as a silencing tactic (Valenti, 2015).

The "boys'-locker-room feel" of the Internet can be seen in largescale patterns of harassment. Although the Internet has been called a democratic space (Haraway, 1991; Sussman & Tyson, 2000), researchers find that not everyone has an equal voice (Megarry, 2014). This is partially because people who are marginalized in other areas of their lives may self-censor their opinions on the Internet to avoid harassment (Fahs & Gohr, 2012; Mantilla, 2013). In contrast, privileged groups are less likely to experience retaliation in online environments: research indicates that white men are the least likely group to be the targets of online harassment, and they are also less likely to interpret negative online experiences as harassment (Lenhart, Ybarra, & Price-Feeney, 2016). This implies that men—especially white, heterosexual men—feel less vulnerable online, which is why they also report low self-censorship when communicating in these spaces.

Privilege and entitlement, not surprisingly, are important components of online harassment. Phillips (2015) argued that those in positions of privilege often have a sense of entitlement to technology and as a result, feel justified in hateful behavior towards marginalized groups. These patterns, for example, are seen in trolling subcultures where young, white men disproportionately comprise the majority of abusers. Unlike other forms of harassment, trolls

make "humorous" or provocative posts with the aim of upsetting someone or eliciting an angry response from them. Much of trolling, Phillips emphasized, is purposely directed at women and people of color to silence their contributions. Phillips located men's motivations for targeting these groups in historical patterns of masculine domination— trolling reflects men's performance of masculinity with the purpose of policing the boundaries of gender in online environments. Accordingly, perpetrators use trolling to communicate the superiority of certain groups under the guise of provocative humor.

Closely tied to Phillips analysis of trolling culture, Mantilla (2015) argued that gender is important when considering online harassment. Mantilla examined narratives of Internet journalists and bloggers to make the compelling argument that women experience *gendertrolling* — a pattern of aggressive and threatening online abuse that is mostly perpetrated by men and that is aimed to publicly shame women over a long duration of time. Although gendertrolls may have individual motivations for harassment, gendertrolling arises from the same misogyny that fuels offline forms of street harassment, sexual harassment in the workplace, and sexual assault. Mantilla contended that gendertrolling upholds a narrow model of masculinity because it validates men's power, status, and dominance over women. Online comments that circulate sexually explicit rhetoric, misogynist epithets, and appearance-related judgments are therefore updated tools used to silence women's contributions (Jane, 2014a).

The past few years have produced noteworthy incidences of harassers who dedicate their time to gendertrolling. For example, feminist activist and journalist Caroline Criado-Perez became a target of online gender harassment after she petitioned the British government to put more female faces on its currency (Hess, 2014). *New York Times* columnist Lindy West, who writes about feminist issues and body positivity, frequently receives online comments that use

sexual violence, appearance-related insults, and gender stereotypes in response to her work (Gross, 2017). Similarly, the *Tropes vs Women* project triggered a campaign of gender harassment against feminist media critic and blogger Anita Sarkeesian. Harassers posted disparaging comments online and vandalized Sarkeesian's article on Wikipedia with sexual images after she called for increased representation of women in video games (Poland, 2016). The thread throughout these incidences is the use of online harassment to mitigate calls for gender equality in traditionally male spaces (Banet-Weiser & Miltner, 2016).

The #Gamergate controversy in particular exemplifies men's attempts to reclaim power, status, and dominance when faced with gender diversification. In August 2014, independent videogame developer Zoe Quinn became the target of an online campaign aimed at defaming her reputation within the gaming industry. The harassment went viral after an ex-boyfriend alleged that Quinn garnered positive reviews for her videogame Depression Quest through romantic involvement with men in games journalism. Although the allegations were false, a group of social media users continued the coordinated harassment campaign against Quinn and other women in the gaming industry under the Twitter hashtag #Gamergate. Supporters of Gamergate ostensibly situated the movement as critiquing ethics in game journalism and with protecting the male "gamer" identity from increasing criticism by feminists and progressives. However, the violent and sexualized vitriol characteristic of Gamergate suggested that the movement's underlying intent was to silence women from participation in on- and off-line public discourse.

Gamergate ultimately functions as a narrative about the role of masculinity in sustaining harassment. As stated earlier, *masculinity failures* may motivate men to act in stereotypically masculine ways under conditions of threat because manhood is socially valued (Berke et al., 2017; Vandello & Bosson, 2013). Men involved in Gamergate occupied a particularly precarious

social position because they were labeled as geeks, often in opposition to "manlier" men (e.g., jocks, preps, bros). Importantly, the harassment characteristic of Gamergate worked to validate dominant constructions of masculinity which society otherwise denies them. Harris O'Malley (2015), a journalist for the *Daily Dot*, contended that for some Gamergate supporters, "women are castrating bitches and whores, non-Gamergate-supporting men are cucks and betas, Internet Tough Guys are lionized, and the dictates of the group must be enforced through harassment." Applying a psychological frame to Gamergate, men who don't fit traditional masculinity norms may strive to achieve masculinity when confronted with real and imagined challenges to their authority in Internet (and gaming) culture.

The fear of women's encroachment in traditionally male spaces make clear a predominant theme: the unprecedented frequency of online harassment coincides with heightened expressions of popular feminism across media outlets (Banet-Weiser & Miltner, 2016). Popular feminism is seen in hashtags like #mencallmethings and #yesallwomen; in websites such as *Jezebel* that link pop-culture with female empowerment; in social media campaigns on Twitter that ostensibly democratize feminist activism (Gill, 2016). While popular feminisms have varied objectives, there is an underlying theme of self-confidence and empowerment for disenfranchised groups (Budgeon, 2011). Women should, for example, "lean in" to leadership positions in the workforce (Sandberg, 2013) and build resiliency when faced with adversity in white, male-dominated arenas (Valenti, 2014). The push for greater recognition of the rights of women, however, is met with what Banet-Weiser and Miltner (2016) referred to as *popular misogyny*— anti-female expression that circulates to audiences in social media. As women continue to advocate for gender equality on- and off-line, some men perceive this as an attack on manhood and respond with silencing strategies.

Consequences for Women Who Speak Out. While most online users in the United States have witnessed harassment, and almost half have been targets, women experience a wider range of online abuse, including more extreme violations (Lenhart et al., 2016). A recent report from the Data & Society Research Institute and the Center for Innovative Public Health Research (2016) found that women were more likely to be harassed or have false rumors spread about them online compared to men. Women were also more likely to feel anxiety, to experience professional consequences, and to take preventative measures—such as suppressing content—to avert future abuse. The frequency and severity of abuse are often compounded for sexual minorities and people of color, meaning that perpetrators may further denigrate women and men on the basis of their race and/or sexual identities (Poland, 2016).

Research in psychology found that the regularity in which women and other marginalized groups experience harassment is on an entirely different scale than white, heterosexual men. For example, in a content analysis of chat room interactions, Meyer and Cukier (2006) established that women's names received an average of twenty-five times more abusive messages than men's names. These patterns were also seen in harassment targeted at people in positions of authority, such as journalists and politicians (Lewis, Rowe, & Wiper, 2017). An unprecedented analysis of over 70 million comments on the *Guardian* website found that women, people of color, and sexual minorities were more likely to be targets of negative comments compared to white men:

> Although the majority of our regular opinion writers are white men, we found that those who experienced the highest levels of abuse and dismissive [comments] were not. The 10 regular writers who got the most abuse were eight women (four white and four non-white) and two black men. Two of the women and one of the men were gay. (Gardiner et al., 2016).

Above and beyond race and sexual orientation, gender of the writer emerged as an important predictor in the severity of comments on the *Guardian* website. Articles written by women received more blocked comments (e.g., abusive) than articles written by men across all news genres. However, women who wrote in male-dominated sections (e.g., sports and technology) received a higher proportion of disruptive comments than women who wrote about stereotypically feminine arenas (e.g., fashion). Even more telling, articles about feminism and sexual assault attracted the highest rates of blocked comments. These patterns signal that while women who are in positions of authority are frequent targets of harassment, women who write in male-dominated arenas or advocate for gender equality often experience *amplified* abuse online.

Women's voices in online environments may carry less authoritative weight because they are seen as a challenge to the status quo. As women gain representation, they may experience backlash in the form of a "swift, aggressive male response to women claiming a space and speaking up on the internet" (Adam, 2005, p. 115). Amanda Hess, a writer at the *New York Times*, emphasized that women who contribute to public debates about gender equality and feminism are particularly vulnerable to harassment on the Internet:

> When I speak with other women who talk about women's issues, whether it's, you know, from abortion to dating, there will...be people [men] who sort of use gender harassment to lash out against people who are specifically taking on misogyny or discussing sexuality. (Mantilla, 2015, p. 32)

As seen in Hess' experiences, women who advocate for political, economic, and social gains often experience reprisals. Some women experience punitive backlash when enacting authority because they are stereotyped as less competent and ambitious than men (Conley, Ziegler, & Moors, 2013; Rudman et al., 2012). Given that competence and ambition are

associated with men, it should incur penalties for women who enact authority because they are breaking traditional gender rules (Rudman et al., 2012). In other words, some men respond with social and economic penalties when confronted with women who violate traditional gender roles.

Thus far, much of the research on online gender harassment has focused on known antecedents and outcomes. However, there has been little attention to the relationship between technology and gender. In the following section, I examine: (1) why anonymous individuals in computer-mediated environments behave differently than identified individuals; (2) how anonymity impacts the scope and frequency of gender harassment.

Theoretical Approaches to Sociotechnical Affordances

Many people lament that there is something inherent in the Internet as a new technology that results in the harassment of women. There are a number of aspects, such as anonymity and limited authority online, that can amplify harassment in digital spaces. However, these effects are influenced by, and can only be understood through, their interaction with social context:

> The Internet doesn't create the urge to harass women…What it does is it makes harassment more *efficient* and *personal*, all at the same time. A man who like to harass women is limited in physical proximity, time restraints, and legal consideration in the real world. Online, however, a man who enjoys harassing women can attack dozens in a short period of time. He can recruit his friends to make the attacks more intense and has a lot more avenues for attack, going through email, Facebook, Twitter, and blog comments. (Marcotte, 2012)

In other words, gender harassment is anchored in and uniquely shaped by the virtual world; however, motivations to perpetrate online gender harassment reflect patriarchal values deeply embedded in U.S. culture.

Sociotechnical affordances provide a useful framework when evaluating the reciprocal interactions between technology, its users, and its social context. Inspired by Gibson's (1979) work on the psychology of perception, Hutchby (2001) defined sociotechnical affordances as structural features of computer technology that enable or constrain certain actions. Importantly, affordances can impact users' positive and negative experiences with technology (Fox & Tang, 2014; Mao, 2014). For example, research indicates that the inability to observe nonverbal indicators of disapproval online can drive some users to engage in negative behaviors (e.g., discrimination) and show bias (e.g., prejudice and stereotyping) that they are unlikely to exhibit in-person (Udris, 2014). Sociotechnical affordances therefore shape online behavior due to the unique possibilities they offer users (Suler, 2004).

One affordance of computer-mediated communication is anonymity, defined as being unnamed or unidentified. Within the context of abusive behavior and harassment online, anonymity can encourage the use of derogatory speech because there is often little "offline" recourse tied to the perpetrator's identity (Khan, Spencer, & Glaser, 2013; Lumsden & Morgan, 2017). The anonymous user is therefore enabled to engage in anti-normative behaviors online such as trolling and other forms of online harassment (Hardaker, 2010). Two approaches from communication studies and social psychology provide context to positive and negative effects of anonymity—social identity model of deindividuation effects (SIDE) and online disinhibition (Lea & Spears, 1991; Suler, 2004).

SIDE Model. The SIDE model proposes that anonymity effects are influenced by, and can be only understood through, their interaction with social context and group-based norms. Based on social identity theory and self-categorization theory, SIDE posits that social identities are central to group polarization (i.e., the tendency for a group to make decisions that are more

extreme than the beliefs of individual members). Anonymous online users experience deindividuation, or a loss of a sense of self, and are more likely to behave according to the particular norms set by the group (Lea & Spears, 1991; Reicher, Spears, & Postmes, 1995). Accordingly, behavioral conformity depends on the social identities salient to the individual within online contexts.

Anonymity has cognitive dimensions that impact the activation of social identities online. When social identity is salient, individuals may depersonalize perceptions of others and the self — that is, they perceive themselves as being part of a social group rather than as an individual (Postmes, Spears, & Lea, 1998). Depersonalization increases the likelihood that individuals will infer identities of others (i.e., outgroup) based on stereotypes, leading to the stereotyping of outgroup members. For example, research has found that within anonymous online environments, participants engaged in depersonalization by identifying themselves with salient social groups, such as nationality, status in education, or gender (Postmes & Spears, 2002). This finding suggests that anonymity may diminish users' identification with their personal identities and increase their conformity with salient social identities.

Anonymity also has strategic outcomes. With the advent of the Internet, scholars posited that anonymous communication would obscure social cues (e.g., gender, race, and class) and equalize online participation (Dubrovsky, Kiesler, & Sethna, 1991). However, research indicates that "offline" hierarchies between groups extend to anonymous online communication (Fox, Cruz, & Lee, 2015). For example, recent survey research on sexual harassment in online video games found that men used gender-based and sexual harassment largely in anonymous environments (Fox and Tang, 2014, Fox and Tang, 2015, Kuznekoff and Rose, 2013). Falling in

line with social identity theory (Tajfel & Turner, 1986), when an ingroup has more power than the outgroup, they may engage in anti-normative behaviors online.

The SIDE model addresses the psychological effects of group polarization in computer-mediated environments. However, it's important to highlight that it has difficulty explaining general *disinhibition*, or the lack of restraint one feels when communicating in non-anonymous online environments (Suler, 2004). Certainly, the SIDE model is right in predicting that social identities and social contexts shape behaviors in anonymous online environments. Given that negative behaviors also occur in non-anonymous environments, there may be additional psychological factors above and beyond salience of social identity.

Online Disinhibition. Suler (2004) developed the *online disinhibition effect* to explain why online users loosen "normative and social constraints of behavior" that are otherwise unacceptable in the offline world. While the online disinhibition effect is not inherently negative, it can manifest as *toxic disinhibition*, which includes activities such as gendertrolling, flaming, and cyberbullying (Dillon & Bushman, 2015; Udris, 2014). Suler (2004) argued that users engage in these behaviors because there are often no "real world" consequences for the offender in online environments. For example, due to the absence of an authority figure to monitor prejudicial actions, some men share sexist comments via Twitter because there is no meaningful reprisal (Fox et al., 2015). The relationship between gender-related attitudes and Twitter are an important part of this picture: men's pre-existing attitudes toward women shaped tweeting patterns, and the structural features of Twitter (e.g., ability to be anonymous) amplified the rate of hostile tweets.

A principal component of online disinhibition is *dissociative anonymity*, or the ability to distance oneself from online behaviors. Many online environments, such as Twitter, do not

require identifiable information. Suler (2004) contended that this lack of connection to identity affords users to feel less vulnerable about engaging in hostile behaviors. Indeed, qualitative and experimental research has found support for the relationship between anonymity and hostile behaviors. For example, Rowe (2015) found that comments posted anonymously on the *Washington Post* website communicated higher incivility compared to identified comments posted on Facebook. Similarly, in an online dilemma task, anonymous participants sent more threats to an experimental confederate than identified participants (Lapidot-Lefler & Barak, 2012). Taken together, these findings indicate that anonymity influences user's disruptive behaviors, or certainly to a greater extent, than offline contexts.

The relationship between anonymity and online harassment has been the focus of recent scholarship. For example, in research exploring motivations for hostile behaviors, perceptions of anonymity predicted men's intentions to engage in online sexual harassment (Ritter, 2014). Online disinhibition effects may drive these behaviors, as people tend to behave less defensively and act more naturally online. In other words, online anonymity is akin to wearing a mask because it allows users to reduce the use of existing social norms that regulate the expression of gender harassment in offline contexts. Because online gender harassment is an extension of real-world behavior, it may therefore be easier for men who experience masculinity threat to lash out at women speaking out on the Internet.

Given the lack of research on the relationship between sociotechnical affordances and gender, the current project explored the roles of anonymity and online disinhibition in amplifying likelihood to gender harass. It would be reasonable to expect that following gender threats, men who engage in anonymous online communication are more likely to post, share, or like online content that expresses gender bias due to a loosening of normative constraints. This raises the

question: how does anonymous online communication amplify men's likelihood to gender harass following masculinity threats?

Current Research

The objective of the present studies is to examine the roles of masculinity, gender identity threat, and anonymity in men's endorsement and perpetration of online gender harassment. In the current research, I investigated this relationship through three empirical studies; see Figure 1 for an overview of the research questions for each study. The goal of this research is to demonstrate that psychological approaches to gender and anonymity are important when examining online harassment.

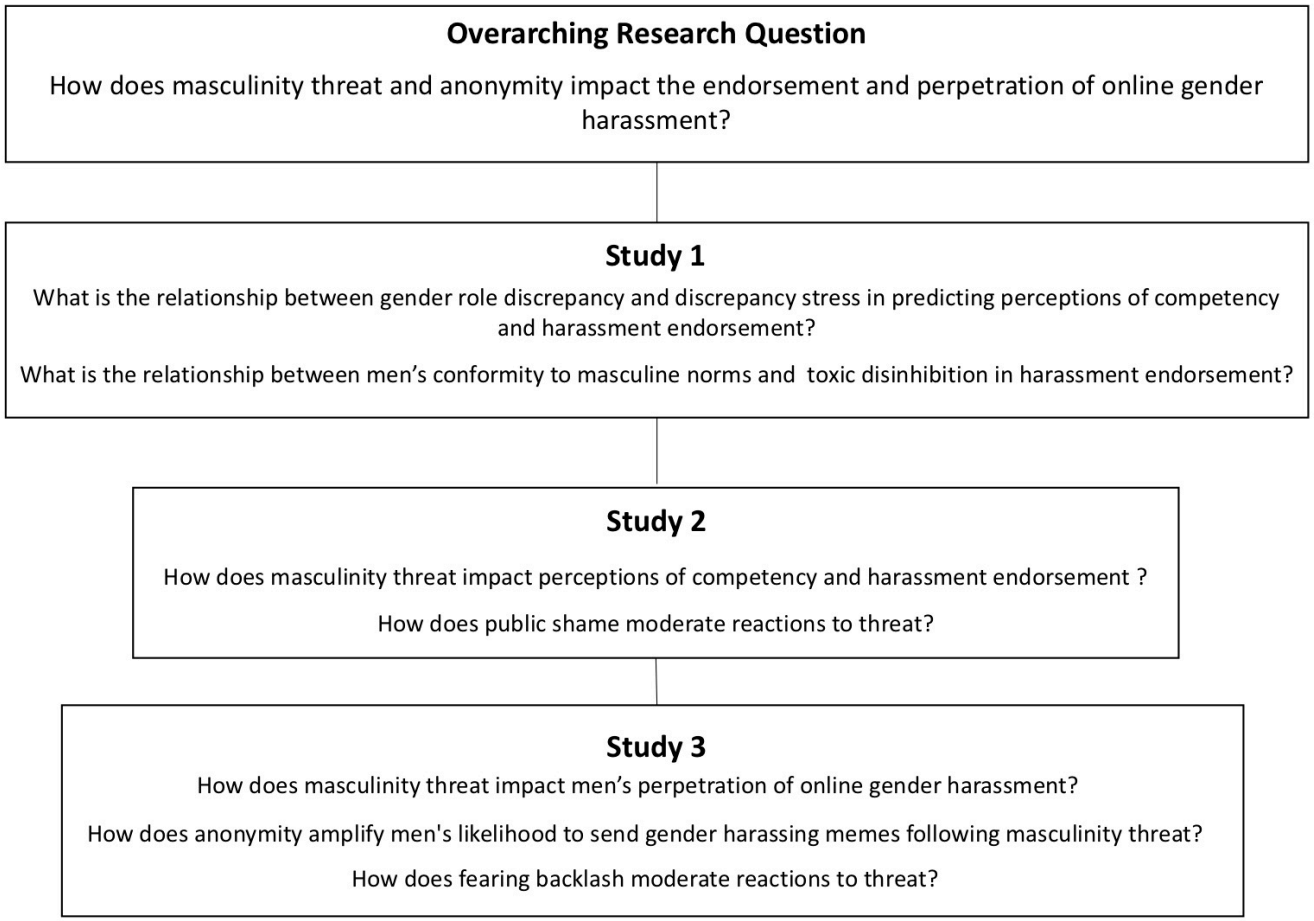
Study 1 focused on the relationship between gender role discrepancy and discrepancy stress in predicting endorsement of gender harassment on Twitter. Given that Study 3 examined the roles of sociotechnical affordances and masculinity threat in men's likelihood to harass, I also investigated the roles of toxic disinhibition and conformity to masculine norms in predicting outcomes of interest. I expect that distress about masculinity failures will be negatively related to competency and positively related to harassment endorsement. Additionally, I expect that conformity to masculine norms and toxic disinhibition will be positively related to harassment endorsement.

The goal of Study 2 was to experimentally examine the relationship between public shame about masculinity failures and harassment endorsement on Twitter following gender identity threat. Assessment of public shame strengthens empirical support of a central tenet of precarious manhood theory—not all men react negatively when they are told that they are less masculine than the "traditional" male. Instead, men who experience *anxiety* about being perceived as insufficiently masculine are likely to engender compensatory behaviors (Vandello & Basson,

2013). I expect that men who experience masculinity threat and public shame about masculinity failures will report lower competency ratings and greater harassment endorsement.

Study 3 joins psychological approaches to masculinity threat with a sociotechnical affordances framework to examine men's *perpetration* of gender harassment on Facebook. Suler's (2004) online disinhibition suggests that anonymity is associated with higher levels of hostility than when individuals are identifiable. However, the relationship between masculinity threat and anonymity have yet to be fully incorporated into existing literature. I expect that men who have their masculinity threatened and who interact anonymously will share more gender disparagement memes compared to men who had their masculinity threatened and were identified. Additionally, I predict that men who experience masculinity threat and fear backlash about their masculinity failures will share more memes. Finally, I expect that the interaction between masculinity threat and anonymity will be stronger for men who report fearing backlash following masculinity threat.

Figure 1. Overview of research questions



CHAPTER 2

Predictors of Competency Ratings and Harassment Endorsement on Twitter

Introduction

Research has found that men's perceptions of their masculinity and the experience of anxiety that may arise from not meeting masculinity norms play an important role in behavior (Berke et al., 2017; Reidy et al., 2016; Reidy, Smith-Darden, Cortina, Kernsmith, & Kernsmith, 2015). That is, when a man views himself as less masculine than the "average" male (*gender role discrepancy*) and believes that others view him to be less masculine as well, he may experience stress due to the self-perception of deficient masculinity (*discrepancy stress*). In these instances, men experiencing masculinity stress may experience social pressure to act in stereotypically masculine ways (e.g., aggression, harassment) because manhood is socially valued (Dahl, Vescio, & Weaver, 2015).

Importantly, men's perceptions of their masculinity can be differentiated from the experience of stress that results from these evaluations. Men who view themselves as less masculine than the "average" male (*high gender role discrepancy*) may be more likely to act out in stereotypical masculine ways given anxiety with being perceived as insufficiently masculine. Alternatively, men who rate themselves as more masculine than the "average" male (*low gender role discrepancy*) may report similar outcomes given anxiety with validating their masculinity to others (see Figure 2 for conceptual model). This raises the questions: do men who experience

high gender role discrepancy stress react in different ways than men who experience low gender role discrepancy stress?

Twitter as a Social Media Context. Study 1 examined Twitter, a social networking service that enables users to track popular (trending) stories on the site. A recent Pew Report positioned Twitter as the "new public square:" a digital space to share information and connect with others through public debate and discussion. Popular press contends that women are able to exert greater social and political influence in this environment (Megarry, 2014) and as a result, "no one [can stop women] from creating [their] tribe and highlighting [their] cause" (Gautam, 2012). Because Twitter allows users to participate in public conversations, it has become an increasingly important platform for women in positions of influence, such as journalists and bloggers, to connect with diverse populations (Parmelee & Bichard, 2012).

Despite Twitter's potential as a democratic space, women are disproportionately targeted for harassment and mistreatment (Citron, 2014). For example, research indicates that female journalists receive roughly three times more negative comments than their male counterparts on Twitter (Barlett, Norrie, Patel, Rumpel, & Wibberley, 2014). When women contribute to public debates that are critical of male dominance, they often contend with heightened digital mistreatment on Twitter (Banet-Weiser & Miltner, 2016). As Amanda Hess (2014) explains, "Twitter is the place where I laugh, whine, work, schmooze, procrastinate, and flirt. It sits in my back pocket... It's become just one of the many online spaces where men come to tell me to get out."

One type of harassment frequent on Twitter is flaming— the posting of inflammatory language, name calling, or insults (Cho & Kwon, 2015). Although flaming occurs in a variety of contexts, it is particularly common when the discussion involves contentious real-world issues

such as women's rights, sexuality, and politics (Hutchens, Cicchirillo, & Hmielowski, 2015). Flaming within these incidences often use gendered insults to derail the conversation and overshadow the discussion of a legitimate topic (Poland, 2016). For example, feminist blogger Anita Sarkeesian received flames such as "back to the kitchen, cunt" (TheDaveKD cited in Sarkeesian, 2012) in response to her advocating for gender equality online. Accordingly, some Twitter users may have the intent of provoking an angry response or argument through flaming.

Hypotheses

My objective for Study 1 was to explore the relationship between gender role discrepancy and discrepancy stress in predicting perceptions of competency and harassment endorsement. I also examined the relationship between toxic disinhibition and conformity to masculinity norms in predicting harassment endorsement. The aims of this study included the following: 1) to determine the role of discrepancy stress in perceptions of competency and harassment endorsement; 2) to explore the interactive effects between gender role discrepancy and stress; and 3) to investigate the relationship between toxic disinhibition and conformity to masculinity norms in predicting harassment endorsement.

Effects of Gender Role Discrepancy and Discrepancy Stress

Given that men who experience *stress* about being perceived as insufficiently masculine may react in ways that reassert masculinity, I expect the following:

- Gender Role Discrepancy Stress will be negatively related to perceptions of competency (Hypothesis 1) and positively related to harassment endorsement (Hypothesis 2).

I predict that stress about being perceived as insufficiently masculine and adherence to norms about manhood plays an important role in men's behaviors:

- I expect a significant interaction between Gender Role Discrepancy and Gender Role Discrepancy Stress. Given that the current study examined whether one's perceptions of masculinity predicted online behavior, I did not have clear a priori hypotheses for the direction of these interactive effects. Men who rate themselves as less masculine than the "average" male may report lower competency ratings and greater harassment endorsement ratings given anxiety associated with masculinity failures. Alternatively, men who rate themselves as more masculine than the "average" male may report similar outcomes given anxiety associated with upholding masculinity (Hypotheses 3 and 4).

Effects of Toxic Disinhibition and Conformity to Masculine Norms

Given that toxic disinhibition may embolden users to engage in behaviors that they otherwise would not face-to-face, I hypothesize that:

- Toxic disinhibition will be positively related to harassment endorsement (Hypothesis 5).

I predict that men's conformity to masculine norms in relationship to toxic disinhibition play an important role in harassment endorsement:

- I expect a significant interaction between conformity to masculine norms and toxic disinhibition. As conformity to masculinity norms increase, I expect that men with more toxic disinhibition will report greater harassment endorsement (Hypothesis 6).

Method

Participants

The online sample included 268 white, heterosexual male participants between the ages of 18-25 who identified as active Twitter users.[1] Participants were recruited and compensated through a paneling service by Qualtrics LCC. Prior to analysis, Qualtrics' staff removed participants who yielded low quality data: those who completed the survey in one-third or less of

the median time (4.5 minutes), those who failed attention checkers (i.e., when participants are asked to indicate a certain response option as a test of their careful reading of each question), and those who provided the same numeric response for all close-ended questions.

The average age of participants was 21.93 years old (*SD*=2.29), and 41% were undergraduate students. In regard to Twitter engagement, participants rated Twitter as an important part of their daily routine (*M*=4.64, *SD*=.96, range=1-6) and reported that they feel part of the Twitter community (*M*=4.26, *SD*=1.28, range=1-6). On average, participants spent 29 minutes per day reading Tweets (*SD*=22.12) and 12 minutes per day posting original Tweets (*SD*=15.59).

Procedure

Participants were asked their age, gender, sexual orientation, and race on the first page of the survey. Participants who identified as white, heterosexual men between the ages of 18-25 were instructed to reflect on the appeal of Twitter content. Participants read the following cover story:

> We are interested in your attitudes about news media on Twitter. You will be shown Twitter profiles of **three journalists**. Demographic information about the journalists (e.g., name) have been anonymized with pseudonyms. Please take a few minutes to evaluate the Twitter profile of each journalist by reading the content. After reviewing each profile, you will be asked a series of questions about your perceptions of the journalist. Additionally, you will be randomly assigned to provide in-depth feedback about **one of the three** Twitter profiles.

To avoid familiarity effects with actual Twitter users, participants were shown fictitious Twitter profiles of three journalists (see appendix A). Mock-ups of social media profiles have

been used in past research assessing attitudes about content on Facebook and Twitter (e.g., Fox et al., 2015; Funk & Coker, 2016). I created fictitious Twitter profiles using *Sketch,* a user interaction design application. With a team of three undergraduate research assistants, I located content from Twitter users who posted about travel, public health, and women's rights. Using this content, I designed three profiles of fictitious Twitter users: Jonathan Miles, a blogger and journalist who writes about travel; Kara Michaels, an author and journalist who writes about public health; and Erin Griffith, an author and journalist who writes about creating an equitable environment for women. Each profile contained eight tweets, and tweet popularity (e.g., likes and retweets) were similar across profiles.

Participants indicated their perceptions of each journalist by viewing static screenshots of their Twitter profiles. Participants rated each Twitter user on several traits (intelligent, trained, expert, informed). Response options ranged from 1(Strongly Disagree) to 6(Strongly Agree), and items were combined to create a composite score for *competency* (α=.92). In order to corroborate the cover story, participants evaluated the Twitter profiles of all three journalists; however, given the aims of the study, the profile of interest was Erin Griffith. Participants were informed that they would provide additional feedback about her profile, including tweets sent to her by Twitter users.

Inflammatory Tweets. Participants evaluated four tweets made in response to Erin Griffith. Undergraduate research assistants located tweets sent to women journalists and bloggers, and I created fictitious tweets (e.g., fictitious username and profile picture) in *Sketch* using this content. One tweet displayed positive feedback, and three tweets displayed inflammatory language that evoked sexist stereotypes (see appendix B). Participants rated each tweet along three dimensions (humorous, agreeable, acceptable), and response options ranged

from 1(Strongly Disagree) to 6(Strongly Agree). Items from the three inflammatory tweets were combined to create an endorsement of harassment score (α=.82).

Pretesting Twitter Profiles and Inflammatory Tweets

To ensure that the Twitter profiles were credible, I asked 82 pilot participants (*N*=23 travel profile, *N*=26 health profile, *N*=33 women's equality profile) to rate each profile on the following dimensions: (a) the degree to which the profile is believable, (b) the degree to which the Twitter user is intelligent, (c) the degree to which the Twitter user is informed. Ratings for believability ranged from 1(not at all) to 4(very believable); ratings for the other two variables ranged from 1(not at all) to 6(extremely).

A repeated-measures one-way analysis of variance (ANOVA) revealed that there were no differences in the degree to which the profiles were believable, F (2, 79) = .637, p=.637. That is, the travel profile (M= 2.35, SD= .75), the health profile (M=2.26, SD= .62) and the women's equality profile (M=2.48, SD= .83) were rated as equally believable. Similarly, the Twitter user for the travel profile (M= 4.19, SD= 1.44), the health profile (M=4.34, SD= 1.19) and the women's equality profile (M=4.0, SD= 1.60) were rated as equally intelligent, F (2, 83) = .168, p=.846. Finally, the Twitter user for the travel profile (M= 3.76, SD= 1.42), the health profile (M=4.13, SD= 1.29) and the women's equality profile (M=3.85, SD= 1.54) were rated as equally informed, F (2, 80) = .421, p=.658.

To ensure that the inflammatory tweets were believable, I asked 33 pilot participants to rate the tweets on the following dimensions: (a) the degree to which the tweets were believable, (b) the degree to which the tweets were funny, (c) the degree to which the tweets were offensive. Ratings for the variables ranged from 1(not at all) to 4(very much). The tweets were rated as

mostly believable (*M*=2.78, *SD*=.79), somewhat funny (*M*=1.70, *SD*=.70), and somewhat offensive (*M*=2.0, *SD*=.66).

Measures

Participants were asked questions about their endorsement of masculinity norms, social media use, and online disinhibition. The presentation of measures, as well as the order of items within each scale, was randomized. See appendix C for survey measures.

Gender Role Discrepancy and Discrepancy Stress. Participants answered five Likert-type questions examining their experience of (1) *perceived gender role discrepancy* (e.g., "Most guys would think that I am not very masculine compared to them," "I am less masculine than the average guy") and five Likert-type questions examining their experience of (2) *discrepancy stress* (e.g., "Sometimes I worry about my masculinity," "I worry that women find me less attractive because I'm not as macho as other guys"; Reidy et al., 2014). Participants rated agreement with each statement using a 6-point scale ranging from 1(disagree strongly) to 6(agree strongly), with higher numbers indicating greater gender role discrepancy and discrepancy stress. Both the gender role discrepancy subscale (α=.94) and discrepancy stress subscale (α=.92) demonstrated good reliability.

Online Disinhibition. Participants responded to an 11-item measure of online disinhibition (Udris, 2014). The *Online Disinhibition Scale* measures disengagement from the social constraints of behavior in online environments through two subscales: benign disinhibition (comfort with openly sharing information in online settings; α = .80) and toxic disinhibition (use of rude language and threats in online settings; α = .79). Example items include: "It is easier to connect with others through ICTs (information and communication technologies) than talking in person (benign disinhibition)" and "It is easy to write insulting things online because there are no

repercussions" (toxic disinhibition). For the current study, only the toxic disinhibition subscale was used as I was interested in the relationship between hostile behaviors and gender harassment. Participants rated agreement with each statement, ranging from 1(disagree strongly) to 6(agree strongly). Higher numbers indicate greater toxic disinhibition.

Conformity to Masculine Norms Inventory (CMNI; Short Form). Participants responded to an 11-item abbreviated version of the *Conformity to Masculine Norms Inventory* (Parent & Moradi, 2011; $\alpha = .79$). The CMNI examines behaviors and attitudes related to 11 masculine norms (Emotional Control, Dominance, Power over Women, Disdain for Homosexuals, Primacy of Work, Playboy, Risk Taking, Self-reliance, Pursuit of Status, Violence, Winning). Sample items include, "If I could, I would frequently change sexual partners" (Playboy) and "It feels good to be important" (Pursuit of Status). Response options ranged from 1(Strongly Disagree) to 6(Strongly Agree). Higher numbers indicate greater conformity to masculinity norms.

Twitter Intensity Scale. To assess connectedness to Twitter, participants responded to an adapted version of the *Facebook Intensity Scale* (Ellison, Steinfield, & Lampe, 2007). I revised the original 6 items to evaluate people's experiences with the Twitter community ($\alpha = .84$). For example, I changed the item, "Facebook is part of my everyday activity" (original) to "Twitter is part of my everyday activity" (revised). Participants responded on a 6-point scale 1(strongly disagree) to 6(strongly agree). Higher numbers indicate greater Twitter connectedness.

Twitter Media Use. To measure the frequency of Twitter use, participants responded to an adapted version of the *Facebook Time Scale* (Junco, 2012). Specifically, participants reported how much they used Twitter per day (slider scale with minutes per day), as well as how many

minutes on Twitter was spent doing various activities (i.e., positing original tweets, reading tweets, responding to tweets, and retweeting). Higher numbers indicate more time spent engaging in Twitter activities.

Results

The observed means and standard deviations for the key variables are reported in Table 1. In order to test hypotheses, I conducted moderation analyses using PROCESS Macros (Model 1) for SPSS v25. First, in two separate analyses, I tested if Gender Role Discrepancy Stress moderated the relationship between Gender Role Discrepancy and outcome variables: perceptions of competency and harassment endorsement. Additionally, I examined if toxic disinhibition moderated the relationship between conformity to masculinity norms and harassment endorsement. Finally, I explored whether results vary by individual difference factors, including connectedness to Twitter. All variables were centered prior to analyses.

Gender Role Discrepancy & Gender Role Discrepancy Stress

I present results for perceptions of competency and harassment endorsement separately. Details of regression equations are presented in Table 2.

Perceptions of Female Journalist's Competency. Hypothesis 1, that distress about masculinity failures will be negatively related to competency ratings, was supported. The model explained a significant proportion of variance in competency ratings, R^2 =.11, F (3, 264) = 11.24, $p < .001$. As stress about meeting masculine ideals increased, men reported lower ratings of competency (b= -.29, p=.006). This relationship was qualified by a significant two-way interaction (b= -.10, p=.01). Simple slopes analyses revealed that Gender Role Discrepancy was significantly related to competency when Gender Role Discrepancy Stress was one standard deviation below the mean (p=.026) but not when Gender Role Discrepancy Stress was one

standard deviation above the mean (p = .94; Hypothesis 3). As Gender Role Discrepancy increases, men with low discrepancy stress reported significantly greater competency ratings, (b=.25, p=.026, see Figure 3).

Harassment Endorsement. When exploring harassment endorsement, the model explained a significant proportion of the variance, R^2 =.14, F (3, 264) = 14.83, p < .001. Consistent with Hypothesis 2, as Gender Role Discrepancy Stress increased, men reported greater harassment endorsement directed at the target (b= .24, p<.001). This relationship was qualified by a significant two-way interaction (b= .129, p<.001). Simple slopes analyses found that Gender Role Discrepancy was significantly related to endorsement of harassment when Gender Role Discrepancy Stress was one standard deviation below the mean (p<.001) but not when Gender Role Discrepancy Stress one standard deviation above the mean (p =.65; Hypothesis 4). As gender role discrepancy decreases, men with low discrepancy stress reported significantly greater harassment endorsement (see Figure 4).

Online Disinhibition & Conformity to Masculine Norms

The model explained a significant proportion of the variance, R^2 =.26, F (3, 264) = 31.15, p < .001 in harassment endorsement. Hypothesis 5, that toxic disinhibition will be positively associated with harassment endorsement, was supported (b= .35, p<.001). Hypothesis 6, that conformity to masculine norms and toxic disinhibition will be positively associated with harassment endorsement, was also supported (b= .14, p=.005). Simple slopes revealed that conformity to masculinity norms was significantly related to harassment endorsement when toxic disinhibition was one standard deviation above the mean (p=.004) but not when toxic disinhibition was one standard deviation below the mean (p=.56). As conformity to masculinity

norms increases, men who reported more toxic disinhibition reported significantly greater harassment endorsement, see Figure 5. Details of regression equations are presented in Table 3. [2]

Post-Hoc Analyses: Controlling for Connectedness with the Twitter Community

A similar pattern of results emerged when conducting all analyses with connectedness to Twitter (*Twitter Intensity Scale*) as a covariate. For competency ratings of the target, Gender Role Discrepancy Stress remained significant after controlling for Twitter connectedness. Twitter connectedness was also significantly related to the outcome; specifically, men with greater connectedness to Twitter reported lower perceptions of competency. Additionally, the interaction between Gender Role Discrepancy and Discrepancy Stress was in the same direction (although the results were trending).

For harassment endorsement, the relationship to Gender Role Discrepancy Stress was in the same direction (although the results were trending). Twitter connectedness also emerged as a significant predictor; specifically, men with stronger connectedness to Twitter reported greater endorsement of harassment. The interaction between Gender Role Discrepancy and Discrepancy Stress remained significant after controlling for Twitter connectedness.

Toxic disinhibition remained significant after considering Twitter connectedness. Twitter connectedness was also a significant predictor, such that men with greater connectedness to Twitter reported greater harassment endorsement. Finally, the interaction between conformity to masculinity norms and toxic disinhibition remained significant after controlling for Twitter connectedness. Details of regression equations are presented in Table 4 and 5.

Discussion

Gender role discrepancy and discrepancy stress are important factors in predicting men's online attitudes. Notably, gender role discrepancy alone did not relate to outcomes; only in

conjunction with stress did gender role discrepancy impact perceptions of competency and endorsement of harassment. These findings suggest that gender role discrepancy is not always associated with experiences of stress, and in turn, negative outcomes. Results illustrate this nuance: men who considered themselves to be less masculine (i.e., nonconformist to masculine norms) than the "average" man and who experience low stress about this discrepancy reported *higher* competency ratings than other men in the sample. It is likely that these men do not place high value on being perceived as masculine and as a result, are less likely to act out in stereotypical ways to demonstrate their masculinity.

In contrast, for highly masculine men, low stress about meeting masculinity ideals was associated with *higher* endorsement of harassment. Men who uphold traditional norms of masculinity may engage in behaviors that bolster public perceptions of masculine prototypicality. This finding is consistent with precarious manhood: given that manhood is an achieved status, it must be consistently reasserted through attitudes and behaviors that uphold masculinity (Vandello & Bosson, 2013). For men who place high value on being perceived as masculine, they may be more likely to display stereotyped masculine behavior to avoid social consequences with masculinity failures (Hunt & Gonsalkorale, 2014).

Although simple slopes analyses indicated significant differences for men with low discrepancy stress, important patterns emerged for men with high discrepancy stress. Results found that as gender role discrepancy increases, more discrepancy stress predicted greater harassment endorsement (see Figure 4). This evidence suggests that men who experience high discrepancy stress may be at increased risk for supporting harassment, perhaps in an attempt to demonstrate their masculinity and reduce anxiety (Reidy et al., 2014).

When considering the influence of technology, toxic disinhibition was a significant predictor of harassment endorsement. Previous research supports this finding: toxic disinhibition is associated with the use of insults or ridicule because of the perceived lack of repercussions and/or anonymity in online contexts (Udris, 2014). However, we gain a more nuanced perspective when considering toxic disinhibition in conjunction with conformity to masculine norms. Among men who conformed more to masculinity norms, increases in toxic disinhibition was associated with greater ratings of harassment endorsement. This finding indicates that the relationship between men's conformity to masculine norms and disinhibition may amplify endorsement of gender harassment in Twitter.

Results were similar when conducting analyses with connectedness to Twitter as a covariate. Interestingly, connectedness to Twitter emerged as a significant predictor: men with greater connectedness to Twitter reported lower perceptions of competency and higher ratings of harassment endorsement. These results could be interpreted as showing that social norms in Twitter may influence user's perceptions of acceptable conduct online. Slurs and harassment targeted at women have been rampant on its platform for almost a decade (Jackson, 2017); Twitter's slow response to curbing misogyny may communicate to its user that these behaviors are permissible online. Alternatively, men who hold misogynistic beliefs may seek out Twitter to support people who share their views. Twitter therefore provides a vehicle to communicate misogyny, yet these beliefs are rooted in individual differences.

Given this evidence, Study 2 further investigates the extent to which anxiety about being perceived as insufficiently masculine impacts harassment endorsement. In particular, I examine how anxiety about masculinity failures impacts harassment endorsement following masculinity threats.

Footnotes

[1] To assess active Twitter participation, participants responded to the Twitter Intensity Scale. Participants who indicated that *Strongly Disagree, Disagree, or Somewhat Disagree* to the scale items were not invited to participate in the study via Qualtrics LLC recruitment. On average, participants identified as active Twitter users (M=4.4, SD=.86).
[2] I ran analyses with age as a covariate; age was not significant and did not change the results.

Table 1. Means, standard deviations for scales

Variables	Mean (SD)	Range
Gender Role Discrepancy	2.76 (*SD*=1.32)	1-6
Gender Role Discrepancy Stress	2.60 (*SD*=1.27)	1-6
Conformity to Masculine Norms	3.43 (*SD*=.78)	1-6
Toxic Disinhibition	2.88 (*SD*=1.19)	1-6
Twitter Intensity	4.40 (*SD*=.86)	1-6
Twitter Media Use	38.23 (*SD*=24.41)	1-100
Posting Original Tweets	12.16 (*SD*=15.59)	1-100
Reading Tweets	29.58 (*SD*=22.13)	1-100
Responding to Tweets	11.86 (*SD*=17.35)	1-100
Retweeting	16.50 (*SD*=22.11)	1-100

Note. *SD*=standard deviation; For most constructs, scale scores were computed by averaging all underlying items. Higher values reflect greater levels of that construct.

Table 2. Results of moderation analyses of GRD and GRDS on competency ratings and harassment endorsement

		Competency Ratings				Endorsement of Harassment			
Step	Variable	*B*	*SE B*	β	R^2	*B*	*SE B*	β	R^2
1	Gender Role Discrepancy	.115	.099	.130	.09	-.138	.083	-.186†	.08
	Gender Role Discrepancy Stress	-.370	.102	-.403***		.340	.086	.433***	
2	Gender Role Discrepancy	.119	.098	.135	.11	-.144	.081	-.194†	.14
	Gender Role Discrepancy Stress	-.292	.106	-.318**		.239	.087	.312**	
	GRD X GRDS	-.100	.038	-.175*		.129	.032	.270***	

Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. *p<.05 **p<.01 ***p<.001 † p<.10.

Table 3. Results of moderation analyses of CNMI and toxic disinhibition on harassment endorsement

Step	Variable	*B*	*SE B*	β	R^2
1	Conformity to Masculine Norms	.163	.057	.123	.23
	Toxic Disinhibition***	.360	.059	.406	
2	Conformity to Masculine Norms	.103	.090	.077	.26
	Toxic Disinhibition***	.352	.058	.397	
	CMNI X TD**	.144	.050	.160	

Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. *p<.05 **p<.01 ***p<.001.

Table 4. Results of moderation analyses with connectedness to Twitter as a covariate

		Competency Ratings				Endorsement of Harassment			
Step	Variable	*B*	*SE B*	β	R^2	*B*	*SE B*	β	R^2
1	Twitter Intensity	-.453	.079	-.332***	.11	.312	.073	.252***	.06
2	Twitter Intensity	-.369	.080	-.271***	.15	.236	.075	.191**	.11
	Gender Role Discrepancy	.082	.096	.093		-.085	.089	-.106	
	Gender Role Discrepancy Stress	-.277	.101	-.302**		.257	.094	.309**	
3	Twitter Intensity	-.337	.082	-.247***	.17	.172	.075	.139	.14
	Gender Role Discrepancy	.088	.095	.099		-.096	.087	-.120	
	Gender Role Discrepancy Stress	-.237	.104	-.258*		.178	.095	.214†	
	GRD X GRDS	-.061	.039	-.107†		.122	.035	.226***	

Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. * $p<.05$ **$p<.01$. ***$p<.001$ † $p<.10$.

Table 5. Results of moderation analyses of CMNI and toxic disinhibition on harassment endorsement, with connectedness to Twitter as a covariate

Step	Variable	*B*	*SE B*	β	R^2
1	Twitter Intensity***	.312	.073	.252	.06
2	Twitter Intensity*	.176	.068	.142	.26
	Conformity to Masculine Norms	.123	.089	.093	
	Toxic Disinhibition***	.348	.058	.393	
3	Twitter Intensity*	.139	.070	.112	.27
	Conformity to Masculine Norms	.081	.090	.061	
	Toxic Disinhibition***	.344	.058	.388	
	CMNI X TD**	.121	.051	.134	

Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. *p<.05 **p<.01 ***p<.001.

Figure 2. Conceptual model of the relationship between GRD and GRDS

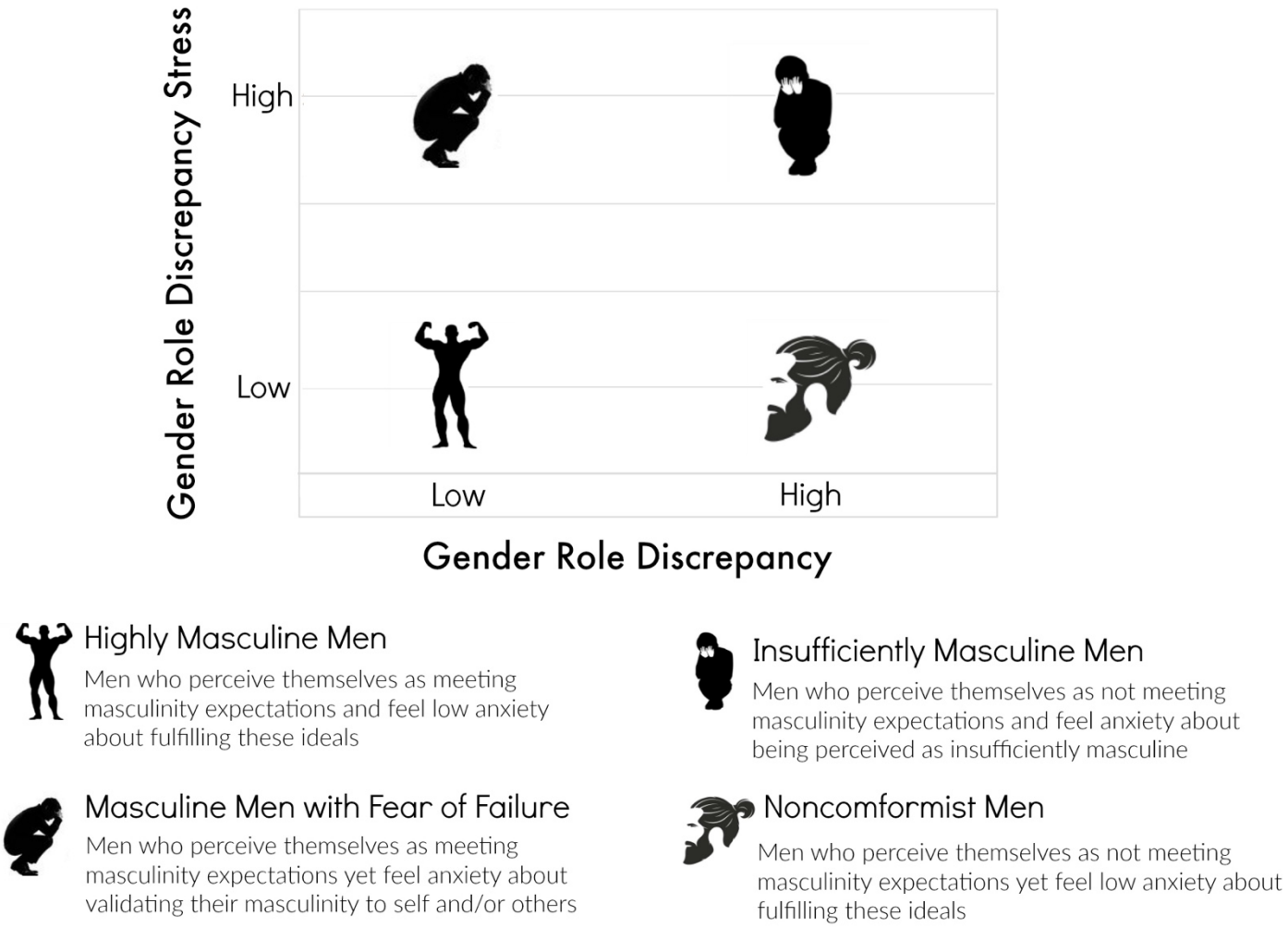
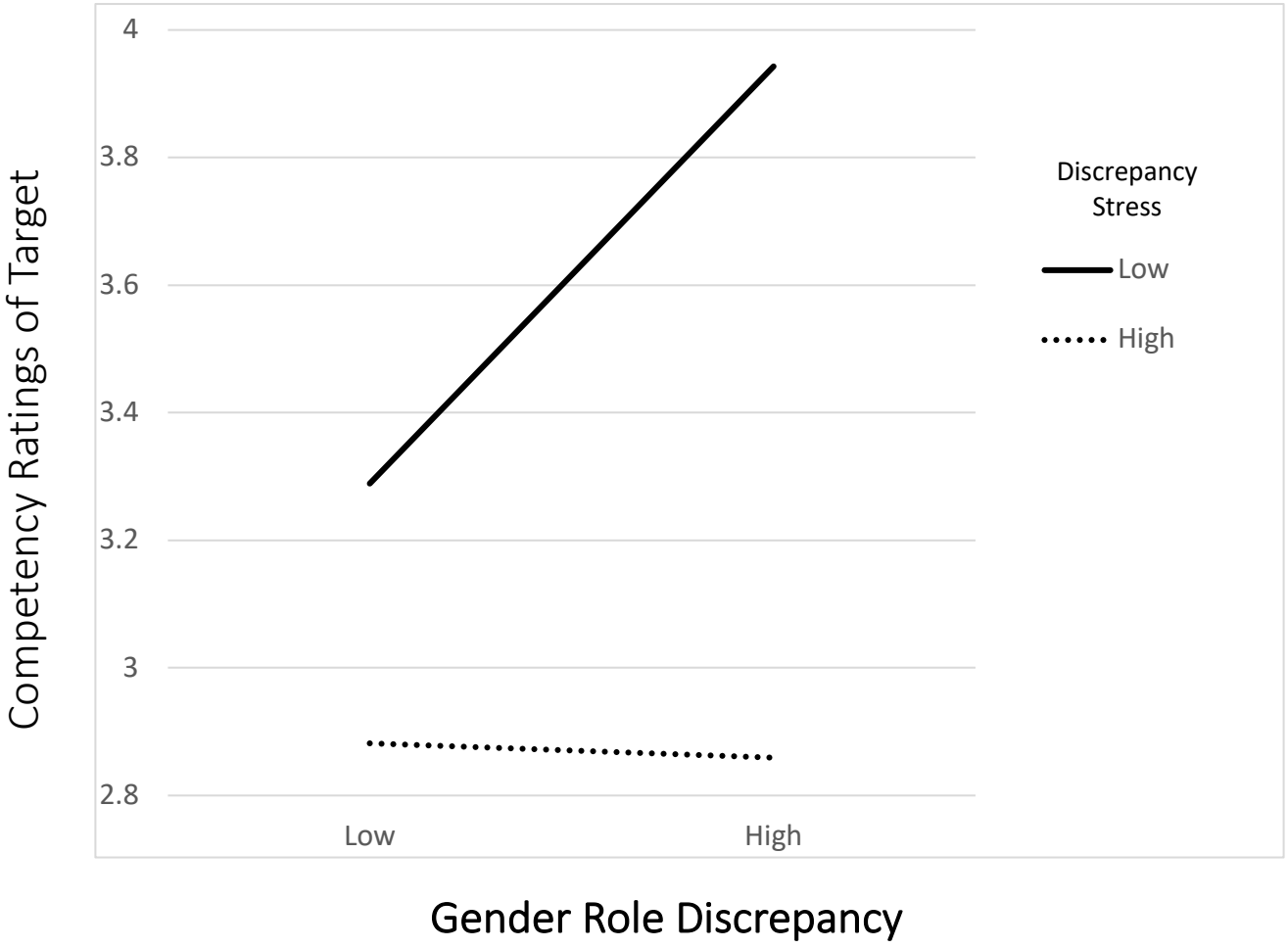
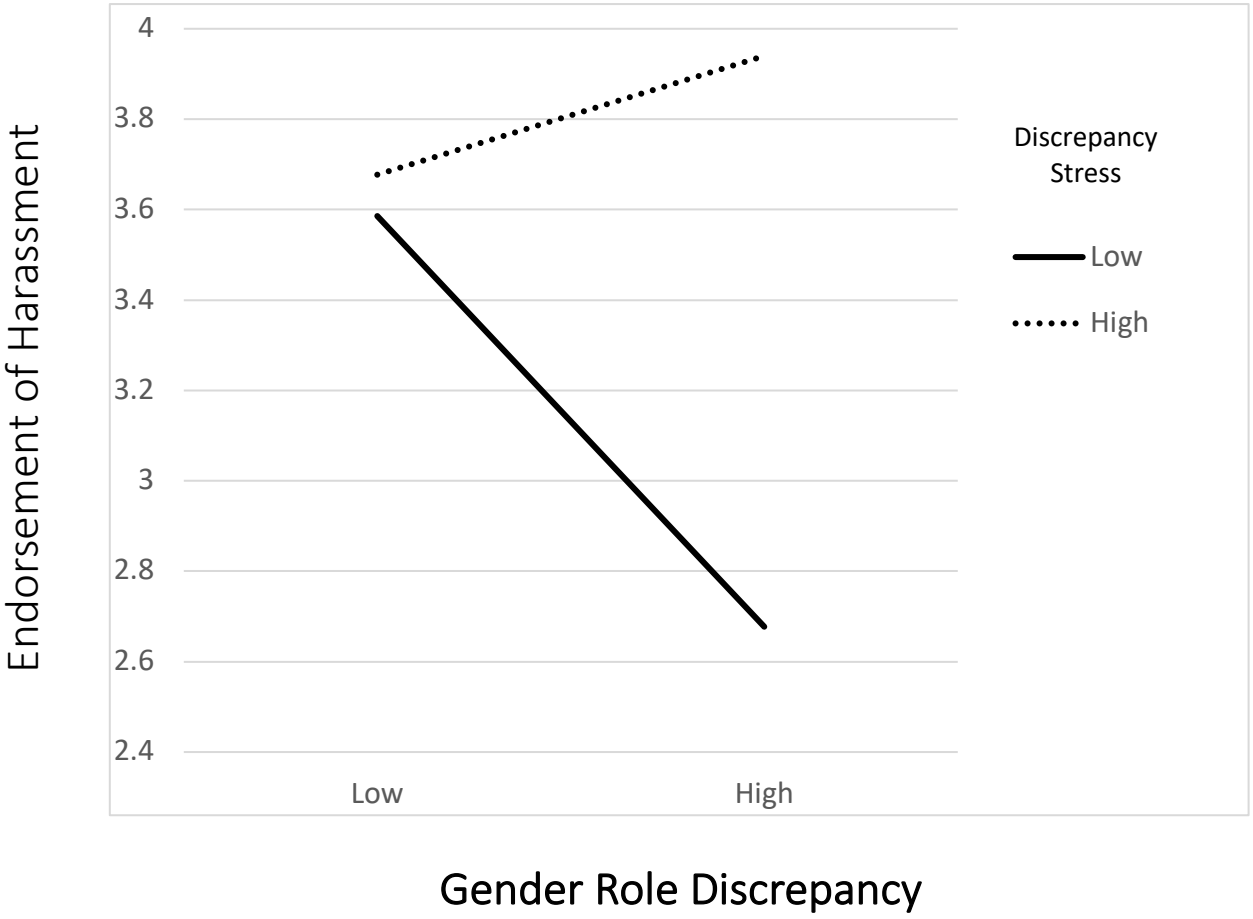


Figure 3. Interaction between GRD and GRDS for competency ratings



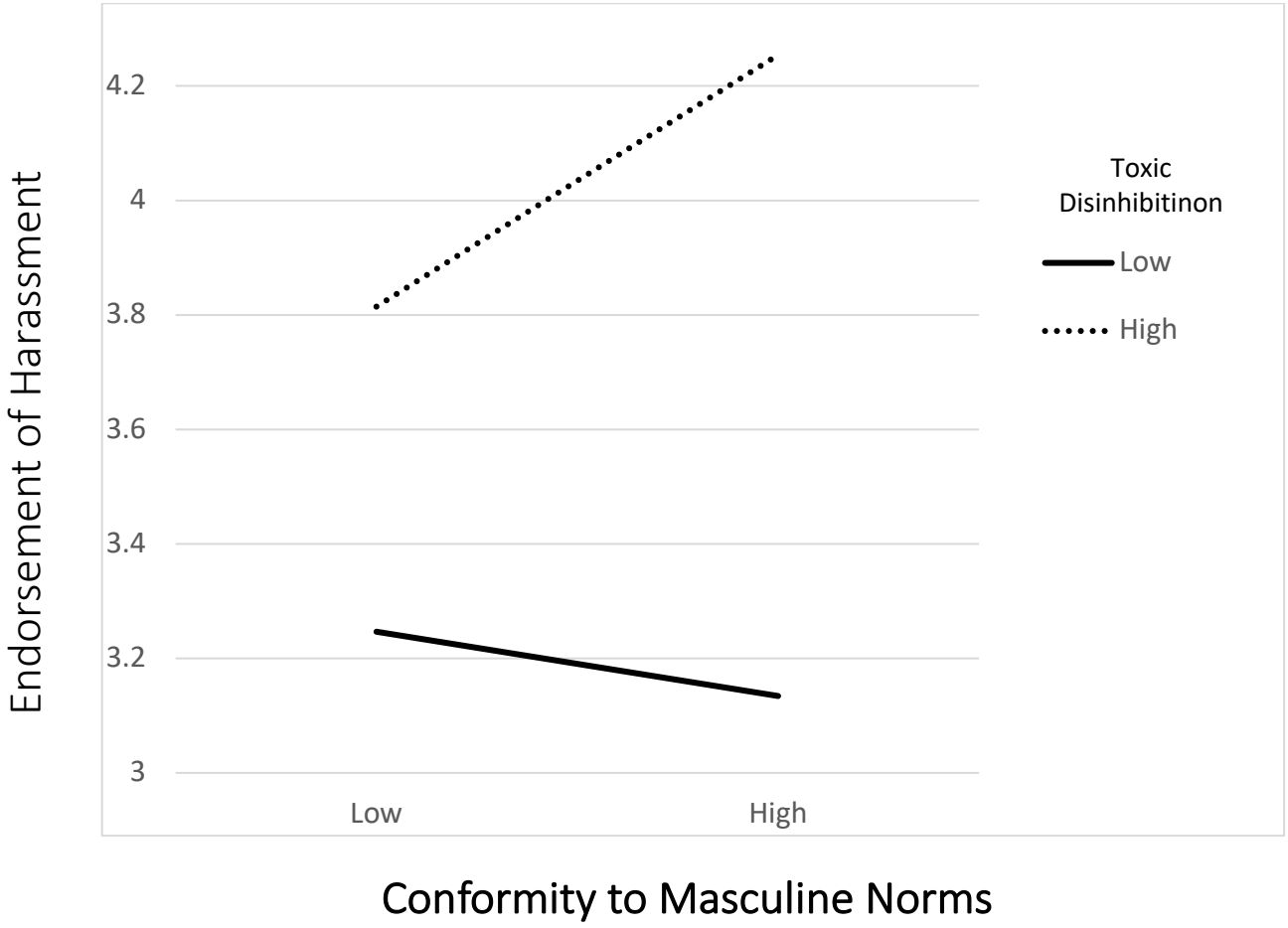
Note. Among men with a high degree for Gender Role Discrepancy, low discrepancy stress was associated with greater competency ratings, (b=.25, p=.026). That is, for men who are nonconformist to masculine norms, low stress about meeting masculine ideals was associated with higher perceptions of competency.

Figure 4. Interaction between GRD and GRDS for harassment endorsement



Note. Among men with a low degree of Gender Role Discrepancy, low discrepancy stress was associated with greater harassment endorsement (b=.24, p<.001). In other words, for highly masculine men, low stress about meeting masculinity ideals was associated with greater harassment endorsement.

Figure 5. Interaction between CMNI and toxic disinhibition for harassment endorsement



Note: Among men who are high conformers to masculinity norms, high toxic disinhibition was associated with greater harassment endorsement, b=.12, p<.01.

CHAPTER 3

Masculinity Threat, Public Shame, and Harassment Endorsement on Twitter

Introduction

Since manhood is a cherished social identity, men may be particularly receptive to situational cues that threaten their masculinity (Moore & Stuart, 2005; Vandello & Bosson, 2013). For example, men who received feedback that they scored below average on a measure of masculinity or who engaged in a traditionally feminine activity were more likely to harass female interaction partners (Maass et al., 2003) and show aggressiveness (Bosson et al., 2009; Vandello et al., 2008; Willer et al., 2013). These results demonstrate that some men compensate for threatened masculinity by espousing attitudes and behaviors consistent with stereotypical masculine norms.

Research has documented an affective threat response that results from men's anxiety about masculinity failures. When faced with masculinity threats, some men experience negative affect and concern that others will perceive them as insufficiently masculine, which I refer to as *public shame* (Bosson & Vandello, 2011; Dahl, Vescio, & Weaver, 2015). Public shame with masculinity failures has been linked to the maintenance of gender-based inequities. For example, Dahl and colleagues (2015) found that insofar as masculinity threat increased anxiety about masculinity failures, men were more likely to endorse or act upon ideologies that implicitly subordinate women. The authors suggest that men may feel more masculine by

asserting power over women, presumably because being more powerful than women is a key component of masculinity. Therefore, negative affect may play an important role in reactions to masculinity threat.

Hypotheses

The goal of Study 2 was to examine the relationship between public shame and harassment endorsement on Twitter following masculinity threat. Study 1 demonstrated that men experiencing distress about their masculinity reported lower competency ratings and greater harassment endorsement directed at women who advocate for gender equality on Twitter. In the current research, I examine how men *respond* to masculinity threats and the role of public shame in men's harassment endorsement. Study 2 has three aims: 1) to determine the role of masculinity threat in competency ratings and harassment endorsement; 2) to explore the interactive effects between masculinity threat and public shame; 3) to investigate the roles of toxic disinhibition and conformity to masculine norms in predicting harassment endorsement.

Effects of Masculinity Threat Condition

Given that men experience social pressure to prove their manhood, they may be receptive to situational cues that threaten their masculinity. I expect the following:

- Men who have their masculinity threatened will perceive the female target as less competent than men who have had their masculinity affirmed (Hypothesis 1).
- Men who have their masculinity threatened will report greater harassment endorsement than men who have had their masculinity affirmed (Hypothesis 2).

Effects of Masculinity Threat Condition and Public Shame

An important theoretical assumption regarding precarious manhood is that men experience social pressure to avoid failures in masculinity (Vandello & Bosson, 2013).

Following gender identity threat, some men may experience public shame due to masculinity failures. I expect the following:

- I expect significant two-way interactions between condition and public shame. Men who experience threat to their masculinity and report more public shame will report lower competency ratings (Hypothesis 3) and greater harassment endorsement (Hypothesis 4).

Effects of Individual Differences

As in Study 1, I hypothesize that toxic disinhibition and conformity to masculine norms play an important role in men's online behaviors. I expect that:

- Toxic disinhibition will be positively related to harassment endorsement (Hypothesis 5).
- I expect a significant interaction between conformity to masculine norms and toxic disinhibition. As conformity to masculinity norms increase, I expect that men with more toxic disinhibition will report greater harassment endorsement (Hypothesis 6).

Method

The experiment consisted of a 2 (presence of a threat to masculinity or presence of an affirmation of masculinity) condition design. Public shame and conformity to masculine norms were examined as continuous individual differences variables.

Participants

White, heterosexual men between the ages of 18-25 participated in a 40-minute laboratory study in exchange for psychology subject pool credit (n=20) or a $15 Visa giftcard ($n$=84). Participants were recruited through two platforms: 1) an online management system that allows students enrolled in psychology courses to participate in psychological research for course credit; and 2) study advertisements posted to University of Michigan undergraduate message

boards via Facebook. Interested participants contacted the research study team and completed a demographic questionnaire prior to participation.

Of the 104 participants, ten were excluded for expressing doubt about the given cover story in open-ended responses. The remaining 94 participants on whom analyses were conducted had an average age of 20.04 years (SD = 1.87, range 18-25 years). When considering Twitter engagement, the majority of participants did not rate Twitter as an important part of their daily routine (*M*=2.53, *SD*=1.76, range=1-6). On average, participants reported that they spent 16 minutes per day actively using Twitter (*SD*=21.30) and 2 minutes per day posting original Tweets (*SD*=4.16).

Materials and Procedure

All study methods were approved by the university Institutional Review Board (IRB). Participants were told that they would be engaging in two studies when they came to lab, and the online survey was divided into separate sections allegedly related to different studies. Participants were given the following cover story:

> In the first study, we will elicit your feedback on a personality measure that we are in the process of developing for an online dating site. You will receive feedback on your scores and will be asked to answer questions regarding your scores, so we can improve the measure for future use. In the second study, you will be asked about your perceptions of news media in Twitter and interactions between Twitter users.

A research assistant instructed participants to complete the personality measure at an individual computer terminal.[3] In actuality, participants completed the Conformity to Masculine Norms Inventory (CMNI; Mahalik et al., 2003), a 46-item inventory that examines adherence and nonadherence to dominant masculine norms (e.g., aggression). Participants were informed

that the CMNI was a "personality test that gives scores along a variety of different personality dimensions." Many of the items in the inventory are not overtly related to gender roles (e.g., "In general, I will do anything to win," "I enjoy taking risks"). During debriefing, most participants indicated that they perceived the test as a general assessment of personality that included some questions related to gender.

False feedback. Based on the prototypicality threat (e.g., Schmitt & Branscombe, 2001) and the masculinity threat paradigms (Hunt & Gonsalkorale, 2014), participants were randomly assigned to receive one of two types of false feedback ostensibly comparing their scores with those of other men who had completed the same personality inventory (*masculinity affirmed, masculinity threatened*).[4] After completing the CMNI, participants were given the following instructions (adapted from Hunt & Gonsalkorale, 2013):

> The research assistant will now examine your scores on the personality test. We have currently identified several different factors, and we will give you your results along three of those dimensions for your feedback. Your feedback is important, so we may make the personality test as accurate as possible for future use. The dimensions you will receive feedback on include: extroversion; openness to new experiences; and masculinity. Your personality scores will be compared to other participants who have taken the measure at University of Michigan.

A research assistant retrieved the personality results from a laboratory in East Hall and delivered printed copies to participants. Participants were shown a scale with an "X" indicating where they scored, alongside an explanation of their results. All participants fell within the average range for the first two (filler) dimensions (i.e., extroversion and openness to new experiences). For those assigned to the *masculinity threatened* condition, the "X" was placed

below the average range. For those assigned to the *masculinity affirmed* condition, the "X" was placed slightly higher than the average range (see Figure 6 for gender identity feedback). Participants reported disappointment and emotional reactions (e.g., anxious, upset, happy) after reviewing their scores.

Twitter profile ratings. After participating in the first task, participants were instructed that the personality test was complete, and they would move on to another study. Participants were told that they would answer questions about their perceptions of news media on Twitter and interactions between Twitter users. As in Study 1, participants evaluated the Twitter profiles of three journalists (see appendix A). They indicated their perceptions of each journalist by viewing static screenshots of their Twitter profiles and rated each Twitter user on several traits (i.e., intelligent, trained, expert, informed). Items were combined to create a composite score for competency (α=.84). Participants also evaluated three inflammatory tweets made in response to Erin Griffith. Items were combined to create an endorsement of harassment score (α=.78).

Finally, participants were asked questions related to their social media use and online disinhibition. Following completion of the study, a research assistant debriefed participants and informed them of the study's purpose. See appendix D for Study 2 survey items.

Measures

Participants completed the same measures of online disinhibition, Twitter media use, and Twitter connectedness as in Study 1. In addition to these measures, I included items related to disappointment and emotional reactions to personality test results. The presentation of the measures, as well as the order of items within each scale, was randomized.[5]

Disappointment with Personality Results. Using 6-point scales 1(strongly disagree) to 6(strongly agree), men reported the extent to which they felt disappointed with their personality

results. Participants were asked, "I am disappointed with my results," and "I feel that my results accurately reflect me as a person." Higher scores indicate greater disappointment with results. Question 2 was used a covariate in analyses.

Public Shame. Men were asked to imagine their personality results being made public. Using semantic differentials, participants answered the following question for each of eight emotions: "Imagine that we publish your personality scores in social media. When you think about your name and scores being published in social media, how _____ do you feel?' (anxious, nervous, defensive, depressed, upset, insecure, sad, angry). Items were averaged across emotions to calculate a public shame score, with higher scores indicating greater public shame (adapted from Dahl, Vescio, & Weaver, 2015, α=.92).

Conformity to Masculine Norms (46-item). The shortened version of the Conformity to Masculine Norm Ideology (Mahalik et al., 2003) is a 46-item inventory that assesses the degree to which participants conform to dominant masculinity norms. In particular, the CMNI examines behaviors and attitudes related to 11 different masculine norms (Emotional Control, Dominance, Power over Women, Disdain for Homosexuals, Primacy of Work, Playboy, Risk Taking, Self-reliance, Pursuit of Status, Violence, Winning). Sample items include, "If I could, I would frequently change sexual partners" (Playboy) and "It feels good to be important" (Pursuit of Status). All questions will be answered on six-point Likert-type scale, ranging from "strongly disagree" to "strongly agree." Items were averaged across emotions to calculate a conformity to masculine norms score, with higher scores indicating greater conformity (α=.88).

Online Disinhibition. The Online Disinhibition Scale (Udris, 2014) was used to assess benign disinhibition (comfort with openly sharing information in online settings; α = .67) and toxic disinhibition (use of rude language and threats in online settings; α = .70).

Twitter Intensity Scale. The Twitter Intensity Scale (adapted items, Ellison, Steinfield, & Lampe, 2007) was used to assess connectedness to Twitter (α = .93).

Twitter Media Use. The Twitter Media Use (adapted items, Junco, 2010) was used to assess frequency of Twitter use.

Results

The observed means and standard deviations for the key variables are reported in Table 6. Table 7 provide means and standard deviations for key variables by condition. For regression analyses, all variables were centered prior to analyses. There were no significant differences in mean CMNI score by condition (all F- values > 1.5, all p -values > .08).

Effect of Condition

To examine if condition (masculinity affirmed= 0, masculinity threatened= 1) affected competency ratings and harassment endorsement, I conducted two independent samples t-tests. I compare the ratings of those in the masculinity affirmed condition to those in the masculinity threat condition. Hypotheses 1 and 2 were not supported. Men who had their masculinity threatened (*M*=4.27, *S*D=.73) did not perceive the female target as less competent than men who had their masculinity affirmed (*M*=4.19, *SD*= .81, *t*(92)=-.50, p=.621). Similarly, men who had their masculinity threatened (*M*=2.81, *S*D=.76) did not report greater endorsement of harassment than men who had their masculinity affirmed (*M*=2.72, *SD*= .92, *t*(92)=-.49, p=.622).

Effect of Condition and Public shame

I conducted moderation analyses using PROCESS Macros (Model 1) for SPSS. I explored if public shame moderated the relationship between condition and two separate outcome variables: perceptions of competency and endorsement of harassment. Perceived

accuracy of results (i.e., "I feel that my results accurately reflect me as a person") was a covariate in analyses.

Hypothesis 3, that men who experienced masculinity threat and public shame about their masculinity failures would report lower competency ratings, was not supported (R^2=.03, F(4, 89) = .72, p=.57). However, consistent with Hypothesis 4, men who experienced threat to their masculinity and public shame about their masculinity failures reported greater harassment endorsement (R^2 =.10, F(4, 89) = 2.52, p < .05). Simple slopes analyses found that condition were significantly related to harassment endorsement when public shame was one standard deviation above the mean (p=.03) but not when public shame was one standard deviation below the mean (p =.21). Among men in the masculinity threat condition, increases in public shame was significantly related to more harassment endorsement (b=.56, p<.02, see Figure 7). In other words, for men who experienced masculinity threat, public shame with masculinity failures was associated with greater harassment endorsement directed at the target. [6] Details of regression equations are presented in Table 8.

Effect of Online Disinhibition and Conformity to Masculine Norms

I conducted moderation analyses using PROCESS Macros (Model 1) for SPSS. The model explained a significant proportion of the variance, R^2 =.23, F(4, 89) = 9.10, p < .001 in harassment endorsement. Hypothesis 5, that high toxic disinhibition will be positively related to harassment endorsement, was supported (b= .31, p<.01). Additionally, conformity to masculine norms was positively related to endorsement of harassment (b= .54, p<.01). Hypothesis 6 – that as conformity to masculinity norms increase, men with more toxic disinhibition will report greater harassment endorsement – was not supported. Details of regression equations are presented in Table 9.

Post-Hoc Analyses: Twitter Connectedness and Conformity to Masculine Norms

A similar pattern of results emerged when conducting analyses with connectedness to Twitter (*Twitter Intensity Scale*) as a covariate. For endorsement of harassment, the interaction between condition and public shame remained significant after controlling for Twitter connectedness. Additionally, the main effects of toxic disinhibition and conformity to masculine norms remained significant after controlling for Twitter connectedness. Twitter connectedness was not a significant predictor of harassment endorsement in both models.

Conformity to Masculine Norms. Research suggests that men high on conformity to masculinity norms are likely to be sensitive to suggestions that they are not meeting expectations of masculine behavior, and thus should react more to a masculinity threat (Hunt & Gonsalkorale, 2014). This raises the questions: Do high conformers to masculine norms behave differently than low conformers to masculine norms within the context of masculinity threat? In order to test this question, I conducted moderation analyses using PROCESS Macros (Model 1) for SPSS. I explored if conformity to masculine norms moderated the relationship between condition and two separate outcome variables: perceptions of competency and endorsement of harassment. Perceived accuracy of results (i.e., "I feel that my results accurately reflect me as a person") was a covariate in analyses.

For competency ratings, conformity to masculine norms was negatively related to competency ratings (b= -.50, p<.01). Men who reported greater conformity to masculine norms indicated lower ratings of competency. Threat condition (b= .05, p=.76) and the interaction were not significant (b= .13, p=.70). For harassment endorsement, conformity to masculine norms was positively related to harassment endorsement (b= .70, p<.001). On average, men who reported more conformity to masculinity norms indicated greater harassment endorsement. Across the two

conditions, participants reported similar endorsement of harassment, (b= .17, p=.33, indicating no main effect of condition). Additionally, the two-way interaction was not significant, b= .13, p=.70).[7]

Discussion

Findings of Study 2 support the idea that public shame about masculinity failures is a key aspect of men's reactions to masculinity threat. Consistent with hypotheses, to the extent that men worried about others' perceptions of their masculinity following threat, they reported greater harassment endorsement. Contrary to hypotheses, and in contrast to prior research that has found an effect of masculinity threat on harassment (Hitlan, Pryor, Hesson-McInnis, & Olson, 2009; Maass et al., 2003), participants reported similar endorsement of harassment across conditions. This highlights that masculinity threat may not increase attitudes consistent with stereotypical masculine norms. Instead, men's negative affect may play an important role in responses to masculinity threat.

Surprisingly, negative affect following gender identity threat did not impact competency ratings. Most prior research links masculinity threat with reparative aggression and dominant behaviors, such as gender harassment and derogation of insufficiently masculine men (Bosson & Vandello, 2011; Cohn et al., 2009; c.f. O'Connor, Ford, & Banos, 2017). Perhaps competency ratings remain unaffected because they do not directly subordinate and devalue women. Alternatively, most men in the sample were undergraduate students, meaning that they regularly interacted with women in a collegiate environment. Men in the sample may view women as competent in a professional setting, yet they may endorse subtle forms of dominance (e.g., endorsement of gender harassment) as a means of repairing threatened masculinity.

Differences in reactions to masculinity threat also emerged. Research assumes that all men react negatively to masculinity threat (Cheryan et al., 2015; Connell & Connell, 2005). However, results indicate that some men reported significantly greater disappointment with their personality test results when they experienced *affirmation* of masculinity compared to when they experienced threat to masculinity. Some men in the sample may eschew dominant masculinity expectations and as a result, report disappointment when their masculinity is affirmed. These findings highlight the importance of considering negative affect in reactions to results—when public shame is considered in relationship to masculinity threat, we gain a more nuanced perspective in harassment endorsement.

Consistent with Study 1, toxic disinhibition was a significant predictor of harassment endorsement. This finding further supports the idea that toxic disinhibition is associated with hostile interactions between persons over the internet (Suler, 2004; Udris, 2014). Contrary to hypotheses, the interaction between conformity to masculinity norms and toxic disinhibition did not replicate. Men's mean scores for conformity to masculine norms were significantly lower in Study 2, meaning that on average the sample is less likely to endorse dominant masculinity norms. Future work is needed to identify the psychological processes in which conformity to masculine norms and disinhibition impact endorsement of gender harassment.

When considering Twitter involvement, men in Study 1 reported significantly higher connectedness and usage. Although users aged 18-25 are more active on Twitter than most (Pew Research Center, 2014), it is likely that the college environment impacts Twitter involvement (Fox, Cruz, & Lee, 2015). College students may use other social networking sites, such as Instagram and Facebook, at different rates than non-college students in their cohort (Knight-McCord et al., 2016). These differences may be particularly important for experimental research,

as college students may be less familiar with Twitter's functionality than other social networking sites (Fox, personal communication, June 29, 2017). I address these concerns in Study 3 by examining Facebook, a social media context familiar to college students.

Given this evidence, Study 3 investigates the role of sociotechnical affordances and masculinity threat in relationship to gender harassment. In particular, I examine how anonymity may amplify men's behavioral likelihood to gender harass following masculinity threat. I also consider the role of fearing backlash for masculinity failures in men's perpetration of gender harassment.

Footnotes

[3] Three women experimenters administered the study.
[4] Before data collection began, 120 personality test results were printed and placed in sealed envelopes (60 affirmed, 60 threatened). This ensured that the experimenters were blind to the feedback condition. A research assistant retrieved an envelope from the lab and delivered the sealed results to participants.
[5] Participants answered questions regarding reactions to the personality test immediately after receiving their results. These measures were not randomized within the survey.
[6] I ran this set of analyses with age as a covariate; age was not significant and did not change the results. The Bonferroni Correction was also used to avoid inflated likelihood of error ($p<.025$). The significance value for Hypothesis 4 was below the threshold ($p=.013$).
[7] I ran analyses examining the relationship between condition, public shame, and conformity to masculine norms. For competency ratings and endorsement of harassment, there was a main effect of conformity to masculine norms ($b= -.49$, $p<.01$, $b= .65$, $p<.001$). All two-way and three-way interactions were non-significant.

Table 6. Means, standard deviations for scales

Variables	Mean (SD)	Range
Conformity to Masculine Norms	3.18 (*SD*=.47)	1-6
Toxic Disinhibition	2.14 (*SD*=.83)	1-6
Twitter Intensity	2.45 (*SD*=1.36)	1-6
Twitter Media Use	16.22 (*SD*=21.31)	1-100
Posting Original Tweets	1.67 (*SD*=4.16)	1-100
Reading Tweets	15.65 (*SD*=18.97)	1-100
Responding to Tweets	1.09 (*SD*=2.41)	1-100
Retweeting	2.89 (*SD*=5.57)	1-100

Note. *SD*=standard deviation; For most constructs, scale scores were computed by averaging all underlying items. Higher values reflect greater levels of that construct.

Table 7. Means and standard deviations by condition

Variables	Affirmation	Threat
Public Shame	2.57 (*SD*=.90)	2.72 (*SD*=1.08)
Disappointment with Results	4.55 (*SD*=.79)	3.58 (*SD*=1.07)
Results Accurate Reflection of Self	3.14 (*SD*=1.0)	3.69 (*SD*=1.16)
Competency Ratings	4.19 (*SD*=.80)	4.27 (*SD*=.73)
Endorsement of Harassment	2.72 (*SD*=.77)	2.81 (*SD*=.93)

Note. *SD*=standard deviation; For most constructs, scale scores were computed by averaging all underlying items. Higher values reflect greater levels of that construct.

Table 8. Results of moderation analyses of condition and public shame on competency ratings and harassment endorsement

		Competency Ratings				Endorsement of Harassment			
Step	Variable	*B*	*SE B*	β	R^2	*B*	*SE B*	β	R^2
1	Results Accurate Reflection	.014	.073	.019	.00	-.055	.288	-.071	.01
2	Results Accurate Reflection	.007	.079	.010	.01	-.104	.086	-.134	.04
	Condition	.076	.166	.050		.122	.180	.072	
	Public Shame	-.011	.085	-.014		.147	.092	.172	
3	Results Accurate Reflection	.011	.078	.016	.03	-.111	.083	-.143	.10
	Condition	.072	.164	.047		.128	.175	.076	
	Public Shame	.140	.125	.180		-.103	.133	-.120	
	Condition X Public Shame	-.266	.164	-.259		.440	.175	.388*	

Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. 0=Affirmation, 1= Threat. *p<.05.

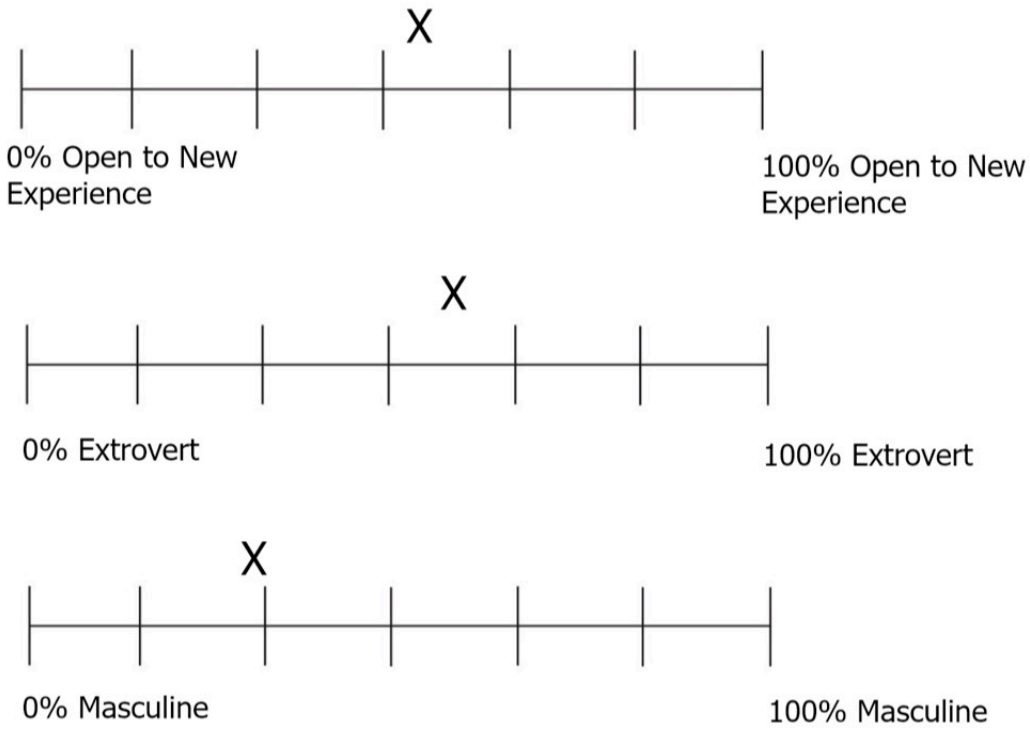
Table 9. Results of moderation analyses of conformity to masculine norms and toxic disinhibition on harassment endorsement

Step	Variable	*B*	*SE B*	β	R^2
1	Conformity to Masculine Norms***	.540	.173	.300	.24
	Toxic Disinhibition***	.313	.098	.406	
2	Conformity to Masculine Norms**	.540	.175	.300	.24
	Toxic Disinhibition**	.313	.099	.307	
	CMNI X TD	-.001	.238	.000	

Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. *p<.05 **p<.01 ***p<.001.

Figure 6. False feedback to personality test

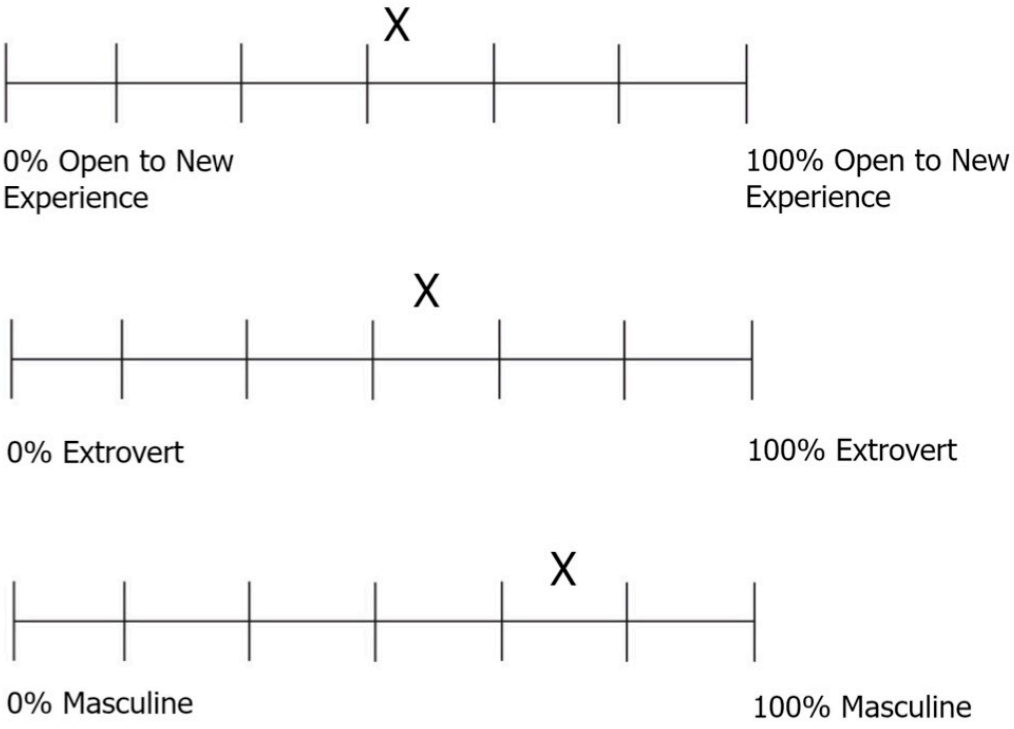
Personality Test Results



Results: The "X" indicates your score on three dimensions of personality compared with other male students at University of Michigan.

- You scored within the **AVERAGE** range for 2 dimensions: extroversion and openness to new experiences.
- You scored within the **BELOW AVERAGE** range on 1 dimension: masculinity. A below average score means that you reported less thoughts, behaviors, and emotions that are traditionally associated with men and masculinity.

Personality Test Results

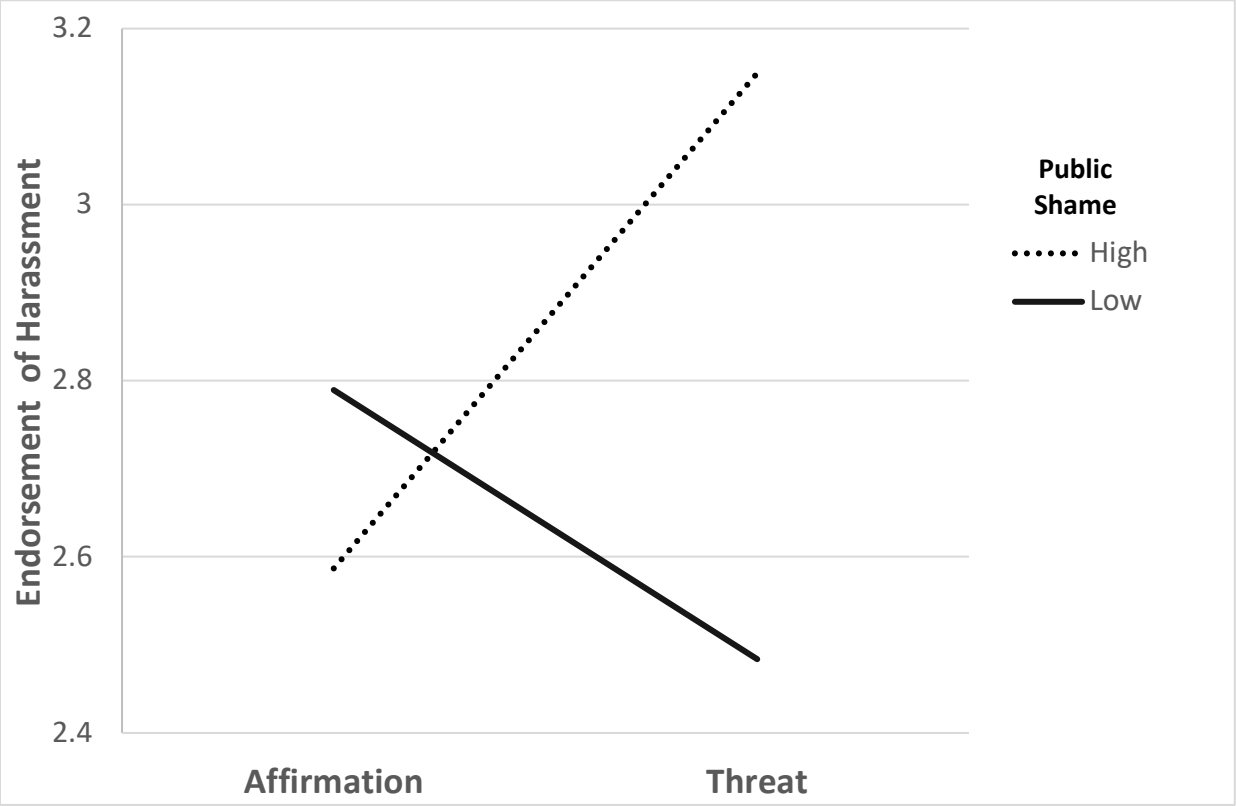


Results: The "X" indicates your score on three dimensions of personality compared with other male students at University of Michigan.

- You scored within the **AVERAGE** range for 2 dimensions: extroversion and openness to new experiences.
- You scored within the **ABOVE AVERAGE** range on 1 dimension: masculinity. An above average score means that you reported more thoughts, behaviors, and emotions that are traditionally associated with men and masculinity.

Note. The first image is the masculinity threatened condition. The second image is the masculinity affirmed condition.

Figure 7. Interaction between condition and pubic shame for harassment endorsement



Note. Among men in the masculinity threat condition, high public shame was associated with greater harassment endorsement (b=.56, p<.02).

CHAPTER 4

Anonymity, Masculinity Threat, and Sharing Memes on Facebook

Introduction

Gender harassment involves a wide range of behaviors including hostile remarks, insults that convey negative attitudes towards women, and sexist humor (Fitzgerald, Drasgow, Hulin, Gelfand, & Magley, 1997). Research indicates that men may be less likely to engage in overt examples of gender harassment (e.g., hostile remarks) due to fear of social opprobrium (Hunt & Gonsalkorale, 2013). Instead, men may engage in ambiguous forms of gender harassment through *disparagement humor* that communicates "the denigration, derogation, or belittlement" of women under the veil of amusement (Ferguson & Ford, 2008, p. 283). Importantly, disparagement humor conveys both an explicit message of derogation of a target, along with an implicit message that the derogation is free of prejudicial motives because it is "just a joke" (Gray & Ford, 2013; Hodson & MacInnis, 2016).

Social media has become an important medium in the creation and distribution of humor about gender through *memes*— a piece of online content which spreads in the form of an image, video, picture, or hashtag (Shifman, 2012). Memes about women in social media can be interpreted as playful humor, yet they can also reproduce misogynistic tropes that subordinate women (Shifman & Lemish, 2010). As one example, popular Internet memes such as "Scumbag Stacy" (who is scantily clad, unethical, selfish and sadistic), "Overly Attached Girlfriend" (who

is an overprotective and clingy girlfriend), and "Cool Girl Carol" (who is sexually adventurous) circulate in websites such as Facebook and Reddit (see Figure 8). Taken together, their content reproduces gender stereotypes about women as unintelligent, hypersexual, or manipulative under the guise of Internet humor.

Although the psychological functions of Internet memes have not been studied extensively in academic scholarship, research in social psychology indicates that men may express amusement with gender disparagement humor as a way to affirm their masculinity (Kehily & Nayak, 1997; O'Connor et al., 2017). For example, O'Connor and colleagues (2017) found that men with greater anxiety expressed amusement with sexist humor following gender identity threat because they believed it would reaffirm their masculinity to others. Absent from the literature is men's *behavioral likelihood* to use gender disparagement humor as a form of harassment when faced with masculinity threats. Given the social permissibility afforded to gender disparagement humor in social media, it is possible that some men use this type of humor to derogate women online and affirm their masculinity.

Importantly, men's use of gender disparagement humor can be understood within the context of anti-feminist bias. As Greenwood and Isbell (2002) argued, amusement reactions to gender disparagement humor varies by the subgroup of women (e.g., feminists, housewives, etc.) that is being targeted. For example, research indicates that men high in hostile sexism sent more sexist jokes to a computer-simulated woman when she was portrayed as holding feminist (as opposed to traditional) beliefs. The authors argued that these findings indicate that harassment, including gender disparagement humor, can be a reaction to gender identity threat (Thomae & Pina, 2015). For women who challenge the gender status quo in social media, they may be

targeted for harassment because they are perceived as a challenge to masculinity (Mantilla, 2015; Poland, 2016).

Backlash Effects for Men. Although both women and men face backlash for violating gender stereotypes, repercussions for men are especially severe (Vandello, Bosson, Cohen, Burnaford, & Weaver, 2008). This is because when a woman violates gender stereotypes, she may be punished or viewed negatively, but her status as a woman is less likely to be called into question in the same way as a man's status (Kimmel, 2008; Vandello & Bosson, 2013). Men who encounter a threat to their masculinity may therefore engage in *behaviors* that bolster manhood to avoid backlash from women and other men (Cheryan et al., 2015).

Study 2 found that public shame played a key role in men's harassment endorsement— to the extent that men worried about others' perceptions of their masculinity following threat, they reported greater harassment endorsement directed at the target. Study 3 builds on these results by examining how fearing backlash impacts men's perpetration of gender harassment on Facebook. Assessment of backlash strengthens empirical support of a central tenet of precarious manhood theory—not all men react negatively when they experience challenges to their manhood. Instead, men who report fearing backlash for masculinity failures may be more likely to engender reparative behaviors (Cheryan et al., 2015; Moss-Racusin et al., 2010)

Facebook as a Social Media Context. Facebook is the most popular social networking site in the United States, with 88% of young adults aged 18-29 reported as using Facebook (Greenwood, Perrin, & Duggan, 2016). Since its establishment in 2004, Facebook has been adopted by a wide range of users who employ the platform for information needs and social interaction (Ellison, Gray, Lampe, & Fiore, 2014). Despite the benefits of Facebook, some users experience harassment and mistreatment that targets their identities (Angwin & Grassegger,

2017). Facebook's support terms partially support this uptick in controversial speech—Facebook's community standards permit material that directly attacks race, ethnicity, nationality, disabilities, religion or sexual orientation if it is satire, humor, or social commentary ("Facebook Community Standards," 2017). Thus, content that constitutes hate speech and harassment may be permitted in Facebook because it is labeled as "just a joke."

College students use of memes in Facebook have recently gained national attention. In June 2017, ten incoming Harvard freshmen had their admissions rescinded after posting sexist and racist memes in a Facebook group message called "Harvard Memes for Horny Bourgeois Teens" and "General Fuckups" (Nathanson, 2017). This group is one of dozens of similar campus Facebook groups, including MIT Memes for Intellectual Beings and Spicy Memelords and Yale Memes for Special Snowflake (Crocker, 2017). Thus, research about gender disparagement memes is timely given the current climate surrounding humor in social media.

Hypotheses

Study 3 examines the interplay between anonymity and reactions to masculinity threat. Study 3 has three aims: 1) to identify if masculinity threat impacts men's likelihood to share gender harassing memes; 2) to identify if anonymity affects men's likelihood to share gender harassing memes, and 3) to clarify the relationship between masculinity threat, anonymity, and fearing backlash on sharing gender harassing memes.

Effect of Masculinity Threat Condition

Men may be particularly receptive to situational cues that threaten their masculinity. I expect the following:

- Men who have their masculinity threatened will share more gender disparagement memes than men who have had their masculinity affirmed (Hypothesis 1).

Effect of Anonymity Condition

Suler's (2004) online disinhibition framework suggests that anonymity is associated with higher levels of hostility than when individuals are identifiable. I expect the following:

- Men in the anonymous condition will share more gender disparagement memes than men who were in the identified condition (Hypothesis 2).

Effects of Masculinity Threat Condition and Anonymity Condition

Given that masculinity threat and anonymity can amplify men's attendant behaviors, a significant two-way interaction will be observed:

- Men who have had their masculinity threatened and who were anonymous will share more gender disparagement memes than men who have their masculinity threatened and were identified (Hypothesis 3).

Effects of Individual Differences

Some men may experience fear of backlash for masculinity failures following gender identity threat:

- I expect a significant two-way interaction between condition and fear of backlash. Men who experience masculinity threat and report fearing backlash will share a greater number of gender disparagement memes (Hypothesis 4).

Effects of Masculinity Threat, Anonymity, and Fear of Backlash

- The interaction between masculinity threat and anonymity will be more pronounced for men who report fearing backlash post masculinity threat feedback (Hypothesis 5).

Method

The experiment consisted of a 2 (presence of a threat to masculinity or presence of an affirmation of masculinity) X 2 (anonymous versus identified) design. Public shame and conformity to masculine norms were examined as continuous individual differences variables.

Participants

White, heterosexual men between the ages of 18-25 participated in a 45-minute laboratory study in exchange for psychology subject pool credit (*n*=216) or a $15 Visa giftcard (*n*=21). As in Study 1, participants were recruited through two platforms: 1) an online management system that allows students enrolled in psychology courses to participate in psychological research for course credit; and 2) study advertisements posted to University of Michigan undergraduate message boards via Facebook. Interested participants contacted the research study team and completed a demographic questionnaire prior to participation.

Of the 237 participants, nineteen were excluded for not meeting recruitment criteria and fifteen were excluded for expressing doubt about the given cover story in open-ended data. The remaining 203 participants had an average age of 18.83 years (*SD* = .96 range 18-22 years). When considering Facebook engagement, many participants rated Facebook as an important part of their daily routine (*M*=3.11, *SD*=1.64, range=1-6). On average, participants reported that they spent 15.21 minutes per day actively using Facebook (*SD*=15.8) and 13.04 minutes per day reading content in Facebook (*SD*=14.36).

Materials and Procedure

All study materials were approved by the university Institutional Review Board (IRB). Participants were given the following cover story:

> We are investigating personality traits of social media users and how they translate humor by sharing memes. Specifically, we are conducting a network analysis tracking how memes get shared and circulate within communities over time. In today's study, you will select content for a public Facebook page called Facebook Thoughts. The followers of this page are students at University of Michigan that we recently solicited to follow the page.[7]

A research assistant informed participants that they would be selecting memes to post to the Facebook Thoughts page from a collection selected by other students who have participated in the study (in actuality, it was a fabricated Facebook page but deception was necessary for the experimental manipulation). In the *anonymous condition*, participants were told that their identities would not be associated with the Facebook page; instead, they would select memes to share anonymously. In the *identified condition*, participants were told that their names and school year would be associated with their posts to the Facebook page. Additionally, research assistants would "tag" the participants in their Facebook posts (i.e., tagging identifies someone else in a post, photo or status update by linking to their Facebook profile).

There are several challenges to conducting research that mimics social media communication. Here, one of those challenges were convincing participants that they were communicating in Facebook to a large audience of University of Michigan students. To create a convincing experience, I created *identified* and *anonymous* Facebook Thoughts pages in *Sketch* with content that mirrored Facebook entertainment pages (see appendix E). I hosted the pages in InVision, a web development application used in prototyping websites and applications. I stripped InVision of its user interaction design functionality through Javascript code (e.g., this hid the tools found in the InVision website and the URL bar of the website). When in full screen

mode, this code allowed the Facebook Thoughts page to appear as if they were *hosted* in Facebook to corroborate the cover story.

Participants were shown a fabricated Facebook Thoughts page to reinforce anonymous or identified accounts. A research assistant scrolled through the page to show participants: 1) the number of University of Michigan followers to the account (N=690) and 2) examples of memes other participants posted to the account. In the anonymous condition, a research assistant showed participants that content was anonymous. In the identified condition, a research assistant showed participants that their name and school year were associated with content. Additionally, the research assistant clicked on a "tagged" post to demonstrate that it linked to a fictitious student's Facebook profile. All content was constant across the Facebook pages except for the anonymous or identified information.

False feedback. After showing participants the Facebook Thoughts page, a research assistant described the study's tasks:

> We are interested in the relationship between personality traits of social media users and humor. In particular, we are interested in how men and women share memes based on unique personality preferences for each group. We determine your personality preferences through two tasks: the first is a personality inventory, and the second is a humor impression task in which you select content for the Facebook page. We will use your responses on the personality inventory and humor impression task to determine the degree to which you prefer male-oriented forms of humor.

Participants were instructed to complete the personality measure at an individual computer terminal. As in Study 2, participants completed the Conformity to Masculine Norms Inventory (CMNI; Mahalik et al., 2003). I manipulated masculinity threat following a similar

procedure as Ford et al. (2017). Specifically, upon completing the personality inventory, participants were presented with the statement: "Scores on the personality inventory range from 0-50, with lower scores indicating less preference for male-oriented forms of humor and higher scores indicating greater preference for male-oriented forms of humor." Participants in the *threat condition* were informed via computer that their score on the personality inventory was 19, which was lower than the average score for men at University of Michigan who have taken the survey. Participants in the *affirmed condition* were informed via computer that their score on the personality inventory was 32, which was close to the average score for men at University of Michigan who have taken the survey (see Figure 9). Participants reported disappointment and fearing backlash after reviewing their scores.

Meme Humor Selection Task. Next, participants completed the humor impression task. Participants were then given the following instructions:

> As part of the humor impression task, you will read several Facebook posts from University of Michigan students. You will be asked to select *one meme* to post in response to their status update. Research assistants will then post some of the memes that you select to the Facebook Thoughts page over the next few days, so we can track its popularity.

Participants read Facebook posts ostensibly from University of Michigan students and were asked to select one of two memes to share in response to their status update. There were 18 Facebook posts in total: 10 posts were neutral in content (written by men and women Facebook users), and 9 posts had a theme of gender equality (written by women Facebook users). For the gender equality content, participants selected from two memes that were either gender disparaging or gender neutral. The number of gender disparagement memes selected was the

dependent variable, with a higher number indicating more gender disparagement memes shared (see appendix F).

Finally, participants were asked questions related to their social media use. Following completion of the study, a research assistant debriefed participants and informed them of the study's purpose.

Pretesting Memes

To ensure that the gender disparagement memes were indeed perceived to communicate sexist humor, 30 pilot participants rated each meme. Following guidelines used by Ford and colleagues (2014), the memes were pretested to the degree to which it disparaged women based on their gender. Ratings for degree of disparagement ranged from 1(not at all) to 8(extremely). The nine gender disparagement memes were rated as significantly more sexist (M = 6.10, SD = 1.23) than the neutral memes (M = 1.92, SD = 1.49), $t(1, 29) = 11.86$, $p < .001$.

To ensure that the gender disparagement memes were perceived equally funny to the neutral memes, 27 pilot participants rated each meme. Ratings for funniness ranged from 1 (not at all) to 8 (extremely). The nine gender disparagement memes were rated as equally funny (M =3.17, SD = 1.491.23) to the neutral memes (M = 3.28, SD = 1.49), $t(1, 26) = .256$, $p < .79$

Measures

Participants completed the same measures of conformity to masculine norms and public shame as in Study 2. In addition to these measures, I included items related to Facebook media use. The presentation of the measures, as well as the order of items within each scale, was randomized.[8] See appendix G for Study 3 measures.

Fearing Backlash. Men were asked to imagine their personality results being made public. They answered ten questions regarding their reactions to the results (Rudman,

unpublished). Sample items include, “Would men give you a hard time (e.g., call you names)?” and “Would you worry about being labeled negatively?” All questions were answered on ten-point Likert-type scale, ranging from “strongly disagree” to “strongly agree." Items were averaged to calculate a fear of backlash score, with higher scores indicating greater fear of backlash (α=.82).

Conformity to Masculine Norms (46-item). The shortened version of the Conformity to Masculine Norm Ideology (Mahalik et al., 2003) examined behaviors and attitudes related to 11 different masculine norms. Items were averaged across emotions to calculate a conformity to masculine norms score, with higher scores indicating greater conformity (α=.86).

Facebook Intensity Scale. The Facebook Intensity Scale (Ellison, Steinfield, & Lampe, 2007) was used to assess connectedness to Facebook (α = .93). This measure includes a series of attitudinal questions that measure the extent to which a participant was emotionally connected to Facebook and the extent to which Facebook was integrated into everyday activities. Items were averaged, with higher scores indicating greater connectedness to Facebook (α=.93).

Facebook Media Use. The Facebook Media Use (adapted items, Junco, 2012) was used to assess frequency of Facebook use.

Results

The observed means and standard deviations for the key variables are reported in Table 10. Table 11 provide means and standard deviations for key variables by condition. For regression analyses, all variables were centered prior to analyses. There were no significant differences in mean CMNI score by masculinity threat condition (F- value > 1, all p -values > .32).

Effects of Masculinity Threat and Anonymity Conditions

I conducted an analysis of variance (ANOVA) to evaluate if sharing gender disparagement memes varied by masculinity threat (condition) and anonymity (condition). I aimed to determine the interactive effects of masculinity threat and anonymity on the number of gender disparagement memes shared on Facebook. Age was a covariate in analyses. Table 12 presents ANOVA results.

Results showed a main effect of threat condition on number of memes shared, such that threatened men (*M*=3.09, *SD*=2.22) shared more gender disparagement memes than affirmed men (*M*=2.44, *SD*= 2.02, $F(1, 202) = 4.87$, $p<.05$; Hypothesis 1). Hypotheses 2, that men in the anonymous condition will share more gender disparagement memes than men who were in the identified condition, was not supported. Men who were identified (M=2.95, SD=2.05) did not share more gender disparagement memes than men who were anonymous (M=2.57, SD= 2.28, t(201)= 1.27, p=.20). Hypothesis 3, that men who have had their masculinity threatened and who were anonymous will share more gender disparagement memes than men who had their masculinity threatened and who were identified, was supported. I found a significant interaction between masculinity threat and anonymity on number of memes shared, $F(1, 199) = 3.52$, $p=.05$). Simple effects tests revealed that threatened men who interacted anonymously (*M*=3.16, *SD*=2.44) shared more disparaging memes than threatened men who interacted identified (*M*=1.98, *SD*=1.84, $F(1, 198) = 9.09$, $p = .003$). There were no differences between affirmed men who interacted identified (*M*=2.96, *SD*=2.0) and affirmed men who interacted anonymous (*M*=2.93, *SD*=2.12, F(1, 198) = .03, p = .863.) Figure 10 presents this interaction.

Effect of Fearing Backlash

I conducted moderation analyses using PROCESS Macros (Model 1). I explored if fearing backlash moderated the relationship between masculinity threat condition and number of memes shared on Facebook. Age was a covariate in analyses. Details of regression equations are presented in Table 13.

Hypothesis 4, that men who experience threat to their masculinity and have more fear of backlash will share a greater number of gender disparagement memes, was supported. Threat condition was marginally significant, such that threat condition predicted number of memes shared (b= .63, p=.058). This relationship was qualified by a significant interaction between masculinity threat condition and fearing backlash about masculinity failures, R^2 =.08, $F(4, 198)$ = 4.27, $p < .05$. Simple slopes analyses found that condition was significantly related to number of memes shared when fearing backlash was one standard deviation above the mean (p=.001), but not when fearing backlash was one standard deviation below the mean (p=.67). Among men in the masculinity threat condition, as fearing backlash increased, they shared a significantly greater number of memes (b= .58, p<.001, see Figure 7). Figure 11 presents this interaction.

Effects of Masculinity Threat, Anonymity, and Fear of Backlash

I conducted an analysis of variance (ANOVA) to evaluate if sharing gender disparagement memes varied by masculinity threat (condition), anonymity (condition), and fearing backlash. Hypothesis 5, that the interaction between masculinity threat and anonymity will be more pronounced for men who report fearing backlash, was not supported, $F(8, 195)$ = .707, p=.69.

Post-Hoc Analyses: Facebook Connectedness and Conformity to Masculine Norms

A similar pattern of results emerged when conducting analyses with connectedness to Facebook as a covariate. For number of memes shared, the interaction between threat condition and fearing backlash remained significant after controlling for Facebook connectedness. Additionally, the interaction between threat condition and anonymity condition remained marginally significant after controlling for Facebook connectedness. Facebook connectedness was not a significant predictor of number of memes shared in both models.

Conformity to Masculine Norms. As in Study 2, I examined whether high conformers to masculine norms behaved differently than low conformers to masculine norms within the context of masculinity threat. I conducted an analysis of variance (ANOVA) to evaluate if sharing gender disparagement memes varied by masculinity threat (condition), anonymity (condition), and conformity to masculine norms.

There was a main effect of conformity to masculinity norms, $F(1, 202) = 10.37$, $p < .001$. Men who reported greater conformity to masculinity norms shared more gender disparagement memes on Facebook. There was also a main effect of threat condition, such that men in the threat condition shared more gender disparagement memes than men in the affirmed condition, $F(1, 202) = 5.15$, $p < .05$. Finally, the interaction between threat condition and anonymous condition remained significant, $F(2, 201) = 4.36$, $p < .05$. Threatened men who interacted anonymously shared more disparaging memes than threatened men who interacted identified.

Additionally, I conducted moderation analyses using PROCESS Macros (Model 1) to evaluate if sharing gender disparagement memes varied by masculinity threat (condition), fearing backlash, and conformity to masculine norms. Conformity to masculine norms was positively related to number of memes shared ($b = .32$, $p < .001$). Threat condition was also significant, such

that men in the threat condition shared more gender disparagement memes than men in the affirmed condition (b= .33, p<.05). Finally, the interaction between threat condition and fearing backlash remained significant, R^2 =.18, $F(4, 198) = 11.07$, $p < .05$.

Discussion

Men who had their masculinity threatened by ostensibly scoring lower on a humor test shared a greater number of gender disparagement memes on Facebook, demonstrating that some men may attempt to appease threats to masculinity through humor that belittles women. This finding supports past research that has found that men initiate and enjoy sexist disparagement humor more when they experience a threat to their masculinity (O'Connor et al., 2017). Fearing backlash also emerged as a key aspect of likelihood to share memes on Facebook. Among men who encountered a threat to their masculinity, men who reported more fearing backlash shared a greater number of gender disparagement memes than men who reported less fearing backlash. Because some men experience social and economic penalties when they break gender norms, they may attempt to counteract this backlash by engaging in compensatory strategies such as conforming more to masculinity norms (Moss-Racusin & Rudman, 2010; Rudman & Fairchild, 2004). Men may initiate gender disparagement humor as a way to reaffirm their masculinity and avoid backlash from women and other men.

Contrary to hypotheses, and in contrast to prior research that has found an effect of anonymity on harassment (Fox et al., 2015; Ritter, 2014), participants in the identified and anonymous conditions shared a similar number of memes. However, significant effects emerged when considering the interactive effects of anonymity and threats to masculinity. Results found that threatened men who interacted anonymously shared more gender disparagement memes than threatened men who interacted identified, and this was a medium effect (d=.53). Perhaps, the

effects of anonymity are amplified when individuals experience masculinity threat because there is no meaningful reprisal for their behaviors in online contexts. Men may attempt to appease threats to masculinity by sharing gender disparagement memes, with anonymity shielding some men from social consequences associated with sexist humor.

Surprisingly, men who had their masculinity threatened and interacted identified shared the least number of gender disparagement memes compared to other men in the sample. Since men often compensate for threatened masculinity through behaviors consistent with masculine norms (e.g., harassment, Hunt & Gonsalkorale, 2013; Maas et al., 2003), it would seem logical that threatened men would respond with some form of backlash against women. Yet Facebook audience characteristics may play an important role in identified men's reactions to masculinity threat. Participants were told that Facebook account followers were University of Michigan students, implying that some of the followers could be acquaintances, classmates, and friends. Perhaps identified threatened men, already experiencing negative affect following gender threat, shared less sexist humor to avoid additional repercussions in their immediate environment.

Research demonstrates that men respond to masculinity threats in a discriminatory manner against feminists (Maas et al. 2003) and women in positions of power (Dahl et al., 2015). When considering social media, gender disparagement memes can provide threatened men a socially acceptable medium to derogate women who challenge the gender status quo (Shifman & Lemish, 2010). Perhaps, some men who experience masculinity threat use sexist Internet humor as a way to demean particular sub-groups of women (e.g., feminists, women in power) who violate traditional gender roles. The use of sexist humor in these situations may serve a self-enhancing function by alleviating one's anxiety about masculinity failures (O'Connor et al., 2017). Importantly, this derogation is not free of prejudicial motives— expressing amusement

with sexist humor can encourage the perpetration of other forms of discrimination against women (Ford, Boxer, Armstrong, & Edel, 2008). Thus, sharing gender disparagement memes in social media is far from being "just a joke."

Footnotes

[7] Followers of the Facebook Thoughts page were ostensibly students at University of Michigan. People can still feel anonymous even when their names are attached to media because the audience does not feel personal to them (Fox, personal communication). For example, people signing in to online comments on news sites with their Facebook accounts doesn't necessarily curtail anonymity given the audience is large and detached. Therefore, even with people's names attached to comments, they may not feel identifiable to the masses in any practical sense. To address this concern, we explicitly told participants that the Facebook Thoughts page was followed by University of Michigan students to lessen anonymity concerns.

[8] Participants answered questions regarding reactions to the personality test immediately after receiving their results. These measures were not randomized within the survey.

Table 10. Means, standard deviations for scales

Variables	Mean (SD)	Range
Conformity to Masculine Norms	3.36 (*SD*=.44)	1-6
Facebook Intensity	2.96 (*SD*=1.3)	1-6
Facebook Media Use	15.21 (*SD*=15.8)	1-100
Sharing Media in Facebook	1.41 (*SD*=3.76)	1-100
Reading Facebook Posts	13.04 (*SD*=13.34)	1-100
Responding to Facebook Posts	1.57 (*SD*=4.0)	1-100
Fearing Backlash	3.6 (*SD*=1.31)	1-10

Note. *SD*=standard deviation; For most constructs, scale scores were computed by averaging all underlying items. Higher values reflect greater levels of that construct.

Table 11. Means and standard deviations for fearing backlash and sharing of memes by threat condition

Variables	Affirmation	Threat
Fearing Backlash	3.34 (*SD*=1.24)	3.86 (*SD*=1.34)
Memes Shared	2.44 (*SD*=2.01)	3.09 (*SD*=2.22)

Note. *SD*=standard deviation; For most constructs, scale scores were computed by averaging all underlying items. Higher values reflect greater levels of that construct.

Table 12. Results of two-way ANOVA for effects of threat condition and anonymity condition on number of memes shared

Variable	*df*	*Mean Square*	*F*	*Sig*	Partial η2
Age	1	21.06	4.83	.029	.024
Threat Condition	1	21.23	4.87	.028	.024
Anonymous Condition	1	6.40	1.47	.227	.007
Threat*Anonymous	1	16.88	3.87	.050	.019
Error	198	4.36			
Total	202				

Note. N= 203

Table 13. Results of moderation analyses of threat condition and fearing backlash on number of memes shared

Step	Variable	*B*	*SE B*	β	R^2
1	Age	-.289	.155	-.130	.02
2	Age*	-.321	.155	-.145	
	Condition*	.625	.303	.146	.05
	Fearing Backlash	.126	.115	.077	
3	Age*	-.320	.153	-.144	.08
	Condition*	.637	.299	.149	
	Fearing Backlash	-.190	.167	-.116	
	Condition X Fearing Backlash**	.586	.221	.261	

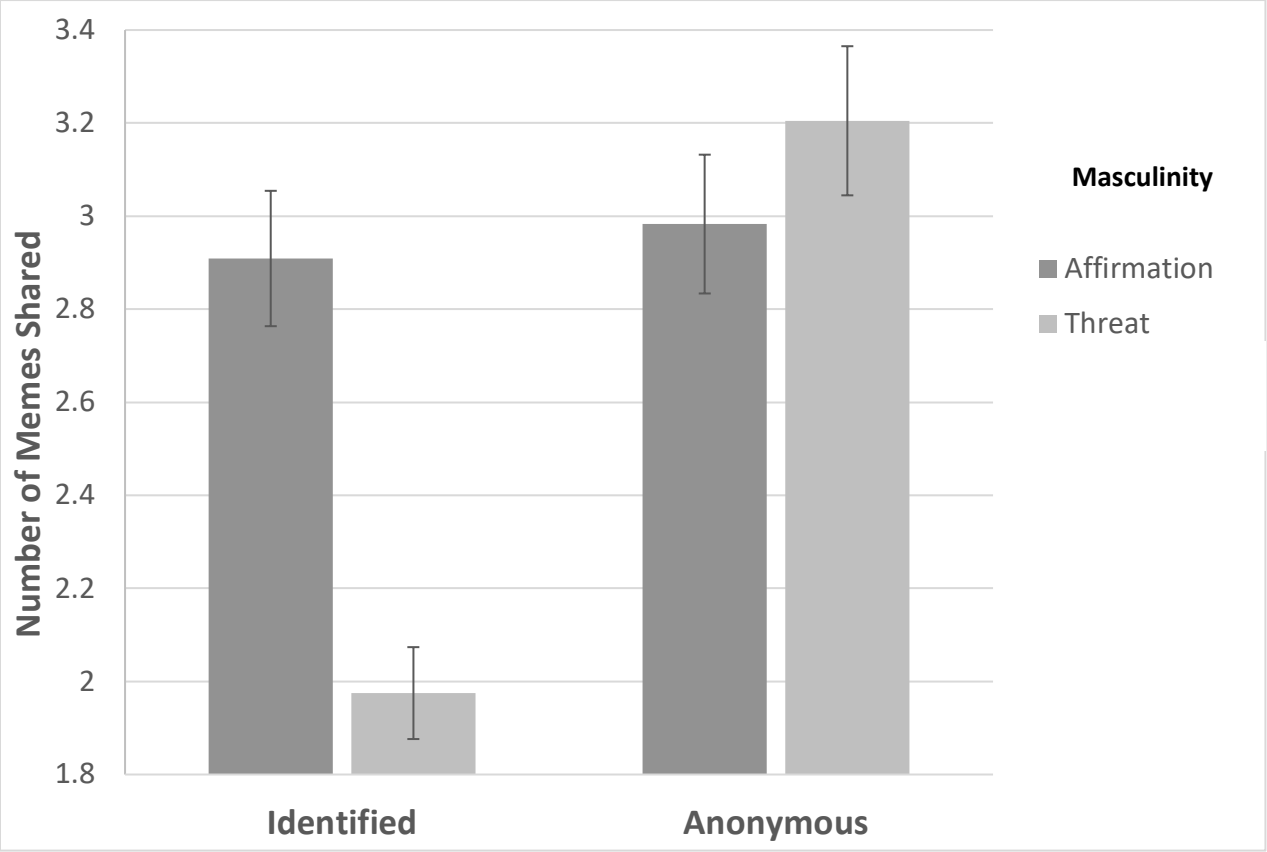
Note. The values of B and β are at step entry. The value of R^2 is cumulative. All variables were centered prior to analysis. *p<.05 **p<.01 ***p<.001.

Figure 8. Examples of gender disparagement memes in social media

Figure 9. False feedback to personality test

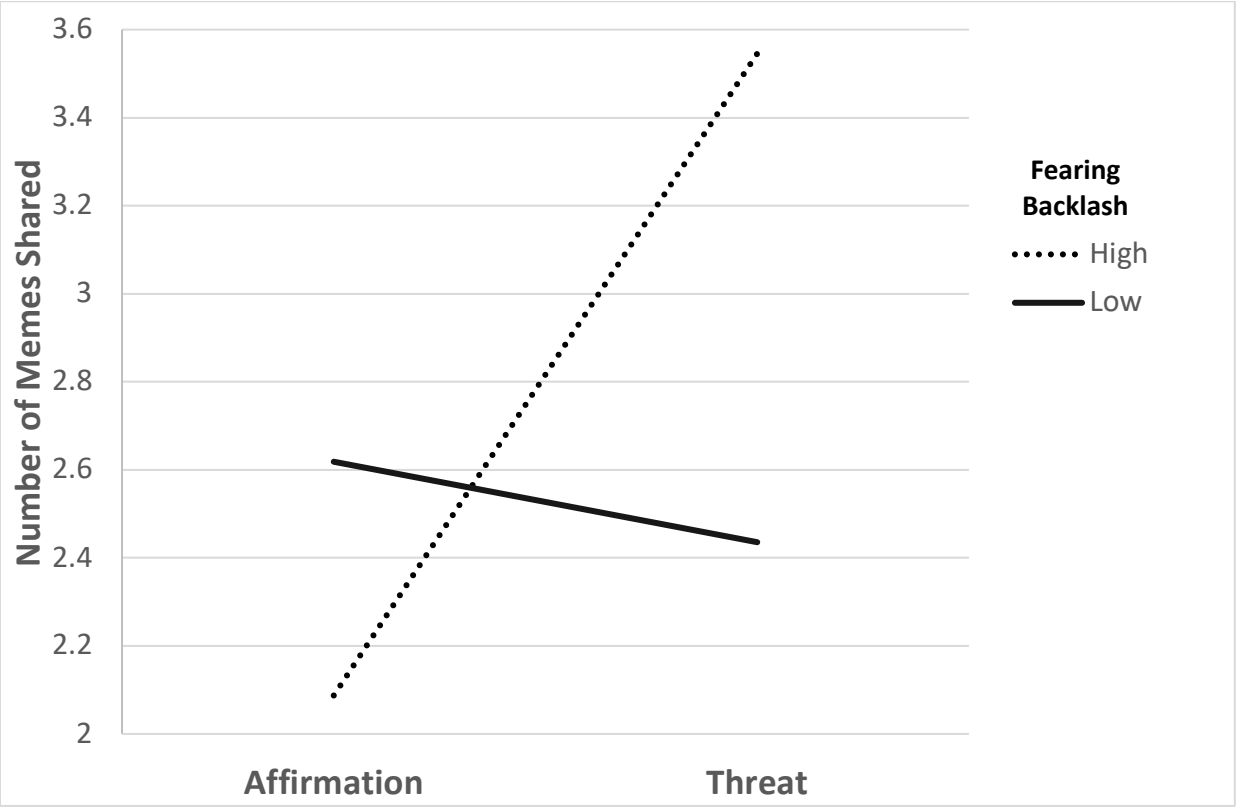
Note. The first image is the masculinity threatened condition. The second image is the masculinity affirmed condition.

Figure 10. Interaction between conditions and memes shared



Note. Error bars represent standard errors. Threatened men who interacted anonymously shared more gender disparaging memes than threatened men who interacted identified, $F(1, 198) = 9.09$, $p < .01$; there were no differences between affirmed men who interacted identified and affirmed men who interacted anonymously, $F(1, 198) = .03$, $p = .863$.

Figure 11. Interaction between threat condition and fearing backlash on memes shared



Note. Among men in the masculinity threat condition, men with high fearing backlash shared a greater number of memes than men with low fearing backlash (b= .58, p<.001).

CHAPTER 5

General Discussion

In the present studies, I examined whether masculinity threat and anonymity impacted men's endorsement and perpetration of gender harassment in social media. Utilizing precarious manhood (Vandello & Bosson, 2013) and online disinhibition frameworks (Suler, 2004), I sought to understand how gender-threatened men's perceptions of being insufficiently masculine affected endorsement/perpetration of gender harassment and how anonymity amplified these effects. In Study 1, I expected that stress about self-perceived masculinity failures would be negatively related to competency ratings and positively related to harassment endorsement. In Study 2, I hypothesized that men who experienced masculinity threat and public shame about masculinity failures would report lower levels of competency ratings and higher levels of harassment endorsement. In Study 3, I expected that men who had their masculinity threatened and who were anonymous would share a greater number of gender disparagement memes than men who had their masculinity threatened and were identified. I also hypothesized that men who experienced masculinity threat and reported fearing backlash would share a greater number of gender disparagement memes.

In support of precarious manhood (Vandello & Bosson, 2013), men's stress about masculinity failures was negatively related to competency ratings and positively related to harassment endorsement in Twitter. However, this relationship was qualified by men's perceptions of masculinity norms— men who considered themselves to be less masculine than

the "average" man and who experienced low stress about this discrepancy reported higher competency ratings and lower harassment endorsement (Study 1). Study 2 found that public shame about masculinity failures played a key role in men's *responses* to masculinity threat. Following masculinity threat, men who expressed concern that others would perceive them as insufficiently masculine reported greater harassment endorsement on Twitter. Study 3 showed that the relationship between masculinity threat and anonymity impacted men's *behaviors* on Facebook. Masculinity-threatened men who were anonymous shared a greater number of gender disparagement memes compared to masculinity-threatened men who interacted identified. Additionally, among masculinity-threatened men, men who reported fearing backlash shared more gender disparagement memes. Consistent with research regarding the roles of gender and sociotechnical affordances in online harassment (Fox et al., 2015; Mantilla, 2013; Poland, 2016), these findings provide evidence that threatened masculinity and anonymity elicits men's support and use of gender harassment in social media.

Masculinity Threat & Online Harassment

Although culture has brought change to gender norms, men still feel immense pressure to uphold manhood through behaviors that demonstrate power, status, and dominance. Unfortunately, there are implications associated with pressure to demonstrate manhood and masculinity. First, some men experience anxiety and stress over conforming to masculine norms, particularly under situations of gender threat (Cheryan et al., 2015; Dahl et al., 2015). Second, some men strive to reestablish their masculinity following gender threat (Bosson et al., 2009; Maass et al., 2003). Third, some men engage in ambiguous reparative strategies (such as sexist humor) to derogate women and reaffirm their masculinity (O'Connor et al., 2017).

Study 1 demonstrated that men's perceptions of their masculinity alone did not relate to outcomes; only in relationship with stress did masculinity evaluations impact competency ratings and harassment endorsement. This nuance is important when considering that some men embrace new gendered practices that allow for more emotionally expressive masculinities (Katz, 2006; Pascoe & Bridges, 2016). Results from Study 1 illustrate variations in masculinity—men who consider themselves to be nonconformist to masculine norms and who experience low stress about this discrepancy report higher competency ratings and lower harassment endorsement than other men in the sample. This evidence suggests that perceptions of oneself as *insufficiently* masculine may impact outcomes only when men experience anxiety about this incongruity.

Studies 2 and 3 demonstrated that some men experience an affective threat response that stems from concern about others' perception of the self. Gender-threatened men who experienced *public shame* (Study 2) or fearing *backlash* from women and other men (Study 3) gave higher harassment endorsement ratings and shared more sexist memes respectively. These findings fall in line with a unifying assumption in the psychology of men and masculinities: manhood must be consistently reasserted through public demonstrations, requiring social validation and recognition by others (Eisler & Skidmore, 1987; Pleck, 1977, 1981, 1995; Vandello & Bosson, 2013). When threatened men face the possibility of having their masculinity failures made public, they may experience anxiety about others' perception of the self and respond with reparative strategies that bolster their masculinity. Perhaps, endorsement and perpetration of gender harassment is an effective strategy to appease public shame and avoid backlash from women and other men.

Interestingly, backlash as a theoretical frame is useful not only in men's reactions to masculinity threat but also men's motivations to perpetrate gender harassment. Results from the

current research demonstrate that some men *fear backlash* when they violate (1) prescriptions that assign men with upholding traits associated with manhood (e.g., toughness and confidence) and (2) proscriptions that men must avoid masculinity failures (e.g., weakness and insecurity; Moss-Racusin et al., 2010). Men who report more fear of backlash for masculinity failures also respond with punitive backlash targeted at women who claim space on the Internet. It seems plausible that threatened men use gender harassment to silence women on the Internet since assertiveness is a prescriptive trait of manhood (but not womanhood). Backlash may therefore bookend narratives of masculinity, as some men respond with punitive backlash to mitigate their own anxieties about backlash for their masculinity failures.

Given that gender is an important frame when considering online harassment, what are consequences for women in social media? It's clear that men's endorsement and perpetration of online gender harassment are located in historical patterns of masculine domination— showing support for inflammatory insults or sharing sexist Internet humor may reflect men's desire to uphold masculinity, especially when they experience anxiety about being perceived as insufficiently masculine by others. With heightened expressions of feminism on Twitter and Facebook, threatened men may use gender harassment to derail conversations and overshadow the discussion of legitimate topics. These tactics have consequences for women: they may be more likely to feel anxiety, to experience professional consequences, and to self-censor as a strategy to avoid future incidences of harassment (Lenhart et al., 2016). Therefore, threatened men may use sexually explicit rhetoric, misogynist comments, and appearance-related remarks as tools to silence women's contributions online.

Whiteness and Aggrieved Entitlement. We live in a particular cultural moment in the United States where white masculinity is highly visible. Raised to expect unrivaled social and

economic privilege, white men navigate what Kimmel (2013) called aggrieved entitlement: a sense that benefits once conferred to white men are being taken away by women, people of color, and immigrants. Aggrieved entitlement is a gendered emotion, an amalgam of threatened manhood and entitlement to reclaim it. Opposition to gains in women's rights, for example, are a mirror to these anxieties—some men employ vitriol and violence toward women who advocate gender equality as a step towards reclaiming their manhood to which they feel entitled. In this sense, aggrieved entitlement can be understood as nostalgia for a time when white men's rightful place in the social hierarchy remained unchallenged. It is, in essence, the possibility for a return to unfettered dominance.

Aggrieved entitlement describes what happens to some men—and groups of men— when systemic changes disrupt gender hierarchies, establishing new opportunities for historically disadvantaged groups. Some men view this shift as a positive development for equal rights and new ways of expressing masculinities. Other men feel excluded by the new cultural landscape and cling to anachronistic masculinity beliefs. Results from Study 1 illustrate this nuance: men who do not place high value on being perceived as masculine and who experience low stress about fulfilling ideals reported the lowest harassment endorsement. In contrast, men who perceive themselves as meeting masculinity expectations and feel high anxiety about fulfilling these ideals reported the highest harassment endorsement. Perhaps, men who place high value on being perceived as masculine are apprehensive about women's growing economic, social and political equality. These men may hold on to old definitions of manhood and the structures that sustained them, resulting in compensatory attitudes that bolster their masculinity.

It's important to note that the current cultural context is one where President Donald Trump posts sexist tweets to demean women. For example, he tweeted that journalist Megyn

Kelly was a bimbo, demeaned journalist Mika Brzezinski for her appearance ("She was bleeding badly from a face-lift") and called Hilary Clinton a variety of sexist names ("crooked Hilary"). These tweets have been shared widely on Twitter and other social media platforms, garnering popularity among his base of supporters. President Trump's derogation of women's gender, appearance, and sexuality demonstrates that sexism is permissible— even celebrated—on Twitter and other social media. These studies were conducted within this context, highlighting the relevance and timeliness of this research.

While this research provides evidence that responses to masculinity threat varies among white, heterosexual men, it is not possible to say whether men of color or sexual minority men would respond differently. Instead, the above research demonstrates that anxiety about being perceived as insufficiently masculine is one possible–and important–component of endorsement and perpetration of online gender harassment among white, heterosexual men. Future research should seek to examine both the mechanisms underlying the responses to masculinity threats among diverse groups of men, and the consequences that follow from them.

Online Disinhibition, Anonymity & Gender Harassment

Perceived affordances of the Internet, such as anonymity and disinhibition, have implications for how people interact in social media. Lending some support to the online disinhibition effect (Suler, 2004), the current research found that toxic disinhibition was a significant predictor of harassment endorsement on Twitter. At times, the relationship between toxic disinhibition and conformity to masculinity norms was important in men's attitudes. Among men who were high conformers to masculinity norms, more toxic disinhibition was associated with greater harassment endorsement. Gender and anonymity were also key aspects in men's online behaviors. Masculinity-threatened men who were anonymous shared a greater

number of sexist memes on Facebook than masculinity-threatened men who were identified. Taken together, these results indicate that elements of online environments may embolden users to support and engage in gender harassment.

It's not surprising that toxic disinhibition predicted harassment endorsement. Considerable research has established that lowering behavioral inhibitions online manifests in aggressive behaviors that would most likely not be exhibited offline (Fox et al., 2015; Fox & Tang, 2014; Lapidot-Lefler & Barak, 2012; Ritter, 2014). This can be seen in online gaming sites (Salter, 2017), cyberbullying (Udris, 2014), comments on YouTube (Moor, Heuvelman, & Verleur, 2010), online relationships (Barak, 2005), and more. What is unique about this research is the relationship between masculinity norm conformity and toxic disinhibition in men's support of gender harassment. Here, there is considerable overlap between negative outcomes associated with toxic disinhibition and masculine stereotyped behaviors. Because there are often no "real world" consequences for the offender in online environments, social media may provide highly masculine men opportunities to assert their masculinity through flaming, harassment, and hostile expressions. It seems plausible that masculinity norm conformity amplifies outcomes associated with toxic disinhibition because highly masculine men are motivated to publicly prove their manhood to others.[9]

The current research also lends some support to the role of anonymity in disinhibited behavior online. Contrary to prior research that has found an effect of anonymity on harassment, participants in the identified and anonymous conditions shared a similar number of memes. Effects only emerged when men experienced masculinity threat: threatened men who were anonymous shared more sexist memes than threatened men who were identified. Falling in line with the online disinhibition effect (Suler, 2004), it's plausible that anonymity amplified the

effects of masculinity threat since it shields men from social consequences associated with sharing sexist humor. In contrast, affirmed men who interact anonymously may share less sexist humor because they are not motivated to prove their masculinity through compensatory behaviors.

The social identity model of deindividuation effects (SIDE) offers an additional explanation for the relationship between gender and anonymity (Lea & Spears, 1991). SIDE posits that individuals experience depersonalization when interacting anonymously online—that is, users perceive themselves as being part of a social group rather than as an individual (Postmes et al., 1998). Because "offline" hierarchies between groups extend online, the interaction of anonymity and salient social identity increases likelihood that individuals will infer identities of others based on stereotypes (Postmes, Spears, & Lea, 1999). As a result, some anonymous users stereotype outgroup members and participate in antisocial behaviors online (Fox et al., 2015; Lea & Spears, 1991).

When considering online gender harassment, men are likely to defer to their masculine social identity when placed in a context where gender identity is activated (Vandello & Bosson, 2013; Weaver & Vescio, 2015). Cues in environments such as Twitter and Facebook further promote the salience of masculinity— for example, aggression targeted at expressions of popular feminisms in social media reinforces core components of masculinity (e.g., power over women; Banet-Weiser & Miltner, 2016). It is likely that the reinforcement of masculine social identity promotes greater depersonalization and stereotyping of women over time (Postmes et al., 1999; Tajfel & Turner, 1986; Thomae & Pina, 2015). Gender harassment is an unfortunate consequence of depersonalization, with some men using flaming and sexist humor as an effective strategy to derogate the outgroup when interacting anonymously. Accordingly, depersonalization

may diminish men's identification with their personal identities, increase their conformity with masculine prototypicality, and increase their derogation of women.

Just as SIDE presents a framework for antisocial behaviors, it also provides theoretical support for enacting meaningful changes in normative pressure (Postmes, Spears, Sakhel, & de Groot, 2001). Research indicates that deindividuation in online contexts increases the awareness of group norms and the responsiveness to cues signaling appropriate behaviors (Postmes et al., 1999). Accordingly, users may be driven to conform to the perceived social norms of a given group and engage in that behavior themselves (Blackwell, Chen, Schoenebeck, & Lampe, 2018). Normative appeals can be a fruitful avenue to reduce gender harassment online— when community guidelines set expectations about how to behave, users are more likely to conform with the social norms of the group (Vikery et al., 2018). This is especially important when considering that platforms such as Twitter have a history of permitting gendered slurs and harassment (Jackson, 2017). Community guidelines that clearly indicate what are unaccepted behaviors from its users may encourage users to avoid those behaviors themselves.

In regard to the experimental design of Study 3, it is important to note that participants in the identified condition were ostensibly "tagged" in the Facebook post. The purpose of tagging participants was to reinforce that their identities would be associated with the content shared to University of Michigan students who followed the Facebook Thoughts page. However, given the functionality of Facebook, tagging a participant would also share this content to an individual's larger Facebook network. Participants in the identified condition may have been mindful of the content they selected since it would be shared with their friends, acquaintances, family, and co-workers. Therefore, it is important to highlight that participant's may have been mindful of two

networks when selecting content in the meme selection task: University of Michigan students who were followers of the Facebook Thoughts page and their larger Facebook networks.

Larger Implications

Given that online gender harassment reflects cultural misogyny, I address larger implications of regulating online harassment. This discussion examines color- and gender-blind approaches to online speech and stakeholders who can limit harassment. By examining the role of social media companies in harassment regulation, I hope to spark a larger conversation about *who is given power to speak* in Twitter, Facebook, and other platforms.

Free Speech & The Role of Social Media Companies

Social media companies have expressed commitment to ensuring that users have the right to express opinions without censorship or restraint, making it more difficult to regulate harassment (Wohn, Fiesler, Hemphill, De Choudhury, & Matias, 2017). At the heart of this argument is a speaker's *intended meaning*—calling someone a "pedophile" can be deemed a verifiable statement, (e.g., criminal charges), but calling someone a “whore” is a matter of personal opinion, and therefore protected under the First Amendment (Marwick & Miller, 2014). These subtle distinctions in a speaker's intended meaning make moderation of harassment woefully vague and highly inconsistent across platforms (Pater, Kim, Mynatt, & Fiesler, 2016).

Tensions between individuals' free speech and the prevention of harm have pushed some social media companies to re-evaluate online harassment and abuse policies. In the wake of public outcry over Gamergate, Twitter has taken steps to monitor harassment: it has established a Trust and Safety Council, prohibited revenge porn, implemented stricter anti-harassment policies, and de-verified users who promote hate speech (Klonick, 2016). Facebook has similarly strengthened its internal policies about online harassment, developed systems for reporting

harassment, and outlined their community guidelines to users (Facebook, 2018). Despite an effort to make these sites safe places for its users, some of these policies have inadvertently benefited majority groups and disadvantaged women and racial minorities (Angwin & Grassegger, 2017).

A 2017 ProPublica report sheds light on Facebook's policy guidelines used to identify hate speech and harassment. For example, one internal Facebook document trains moderators on how to implement the company's global hate speech algorithm. The document asks moderators to identify the group protected from hate speech: female drivers, black children and white men. Falling in line with Facebook's color- and gender- blind approach to divisive speech, the correct answer is white men (Angwin & Grassegger, 2017). The reason for this choice is that Facebook removes attacks directed at *protected categories* based on race, sex, gender identity, religious affiliation, national origin, ethnicity, sexual orientation and disability/disease. Subsets of categories— such as children and drivers— are not protected under these guidelines. White men are considered a group because of their protected identities (race and gender), while female drivers and black children are subsets and therefore not protected.

The irony of this distinction is that it is not in the spirit of equal protection (Citron, 2014). The United States has a longstanding history of permitting preferences for women and racial minorities to redress discrimination (Bernstein, 2015). Facebook’s algorithm protects all genders and races equally, meaning that this approach will "protect the people who least need it and take it away from those who really need it" (Citron, quoted in Angwin & Grassegger, 2017). Color- and gender-blind approaches to online speech have implications for users: Facebook's policies have categorized marginalized groups as unprotected —and therefore free to be targets of divisive speech — while privileged groups as protected from abuse.

Absent from discussion is perhaps a more fundamental question: how does a gender-blind approach to harassment impact women's online participation? Franks (2012) argued that the result of harassment is to position women as a marginalized class, using sexual objectification and gender stereotyping to make women feel unwelcome online. These outcomes demonstrate that just as free speech functions as an individual protection, language can also limit an individuals' right of expression by suppressing speech (Marwick & Miller, 2014). Given the regularity in which women and other marginalized groups experience harassment, minority groups may therefore suppress speech to mitigate harm from majority groups. Divisions between majority and minority users ultimately contribute to a growing digital divide in public discourse: harassment impacts *who is given power to speak* in social media. Strategies to decrease user inequities will only become timelier as social media gains prominence as the new public square.

Stakeholders in Monitoring Harassment

Thus far, technology companies have mostly positioned online speech as a First Amendment issue. However, should we be using the framework of "free speech" to understand online speech? Given that online harassment does not occur in a vacuum and instead reflects the culture at large, it is important to examine institutional practices that shape online misogyny. Next, I will consider three stakeholders who can contribute to cultural change in harassment: social media companies, UX designers and developers, and diversity in technology.

Social Media Companies. Facebook, Twitter, and other platforms have each developed content moderation strategies, algorithms trained to identify violations automatically, and community standards for its users (Duggan, 2017). Although some incidences of harassment are identified and removed, others are not and gain popularity through user engagement (Vikery et al., 2018). The amplification of harassment (e.g., Gamergate) highlights an integral component

of social media business models: platforms hold popularity as a fundamental value, meaning that popularity is a core metric in promoting user engagement and increasing profit through advertising revenue (Bartow, 2009). Harassment takes advantage of popularity by amplifying hate—gendered and racialized harassment generate likes, views, and retweets. When profits drive corporate decision-making, policies that moderate harassment become complex:

> The major Internet platforms are driven by a profit motive. Very often, hate, anxiety, and anger drive participation with the platform. Whatever behavior increases ad revenue will not only be permitted, but encouraged, excepting of course some egregious cases. (Rainie, Anderson, & Albright, 2017)

From a business perspective, harassment and abuse are just as beneficial to platforms as other forms of participation. For example, Reddit has built its brand image through capitalizing on the massive momentum around trolling and harassment that frequent its site. One could argue that Reddit even profited from racist and misogynist abuse by growing its usership, increasing advertisement revenue, and amplifying name recognition (Johnson, 2016; Poland, 2016). Would Reddit increase profits if it actively monitored harassment, or do harassers bring user activity to the platform? Under business models built on popularity, Reddit and other providers will most likely impose rules with the least consequences to grow its usership and increase engagement.

To take harassment seriously, social media companies should develop a different business model that prioritizes its users (Vikery et al., 2018). This approach means that social media companies should take complaints more seriously and drop users who engage in racism, misogyny, and other forms of abuse. Although usership may decrease in the short-term, stringent approaches to harassment prioritizes the long-term goal of encouraging people to engage freely with a platform. Social media companies therefore have a choice: should value be placed on

generating popularity through *all speech*, including harassment? Or should value be placed on developing progressive changes that largely promote protections for its users?

UX Designers and Developers. User experience (UX) designers and developers can also support users who are targets of online harassment. Both UX designers and developers play a key role in the functionality and usability of a website—for example, designers can build tools that allow users to report harassment. These tools include easily accessible online forms for reporting harassment, visible links that connect users with content moderation specialists, or community guidelines about anti-harassment tools available on the platform. Unfortunately, the current approach that platforms like Twitter have taken expects its users to curb harassment (Hudson, 2014). For instance, platforms may expect victims of harassment to detail their experiences in long forms, often with little response about the outcome of the complaint (Jhaver, Ghoshal, Bruckman, & Gilbert, 2018). Platforms may also anticipate that individuals will create tools to mitigate harassment (Vikery et al., 2018). These approaches suggest that harassment is an individual problem rather than a systemic issue.

UX designers and developers should create community guidelines early on in the development process to foster positive user experiences with combating harassment. Community guidelines help cultivate injunctive norms— customary standards for how users *should* behave in an online community (Klonick, 2015). Injunctive norms can be self-enforced by a person's desire to conform to community standards, and they can be externally enforced by the community when an individual violates a norm (Cialdini, 2007). UX designers and developers could help curb harassment by establishing community standards of behavior. Ultimately, by considering the ways that injunctive norms impact diverse audiences, designers and developers

can cultivate spaces where harassment is not normalized (Blackwell, Dimond, Schoenebeck, & Lampe, 2017).

Diversity in Technology. Implementing solutions to harassment may be difficult given the lack of diversity in the technology industry—platforms are run by privileged groups who tend to believe that structural inequities are absent in technology. Not surprisingly, the worldviews of straight, white men powerfully shape communication strategies in social media. Gillespie, a Principal Researcher at Microsoft Research New England, argued that engineers and entrepreneurs excel at building platforms "designed like brutalist economic markets," where users compete for voice and the First Amendment protects silencing tactics of harassers (Vikery et al., 2018). This limited perspective ultimately dismisses the experiences of those who endure structural inequity in their daily lives: women, racial and ethnic minorities, and gay, queer, and transgender people (Banet-Weiser & Miltner, 2016).

If the world is truly fair and meritocratic, workplace culture at social media companies should reflect diversity. However, first-hand accounts detail rampant sexual harassment, gender discrimination, bullying, and racial bias in tech. Findings from the *Tech Leavers Study*, a national study examining why people left their jobs in tech, found that women experienced significantly more unfairness than men (Allison Scott, Kapor, & Onovakpuri, 2017). When taking into account gender and race, underrepresented women and men of color experienced stereotyping at twice the rate of White and Asian women and men (Allison Scott et al., 2017). Gender and racial mistreatment in the workplace ultimately contribute to a toxic environment:

> The culture was toxic. The CEO clearly lacked respect for women. Inappropriate remarks were made about women interviewing for roles in the case that the founder found them attractive. Inappropriate / sexual remarks about women were made in front of his female

> employees during off sites. (There was) tons of micro-management and lack of trust in the abilities of the women who worked for his company. After I left, all of the other women quit too. It was not a female-friendly company. (Latina engineer, quoted in (Allison Scott et al., 2017)

Taken together, these findings raise an important question: If tech and social media companies reproduce structural inequities in the workplace, how do these inequities shape decision-making around the regulation of harassment in online spaces? It is clear from large-scale studies that tech workplace environments are hostile to women and people of color; these biases may also make room for purposefully vague and inconsistently enforced harassment policies in social media.

Increasing diversity of Silicon Valley may expand inclusion, yet meaningful change can only take place when companies take seriously gender and racial biases in the workplace. Interventions include establishing oversight and accountability for how personnel decisions are made, taking seriously reports of harassment and discrimination, and enacting professional consequences for perpetrators (Biebly, 2013). With changes to company culture, teams might build interventions to reduce structural inequities in technology.

Limitations and Future Directions

Although the present set of studies have many strengths, they are not without limitations that future research will want to address. For example, I made my best attempts to conceal the purpose of the research in recruitment advertisements and experimental prompts. However, it is possible that some men guessed the true intentions of the research and biased their answers. This may be especially relevant for men who were targeted for recruitment through poster advertisements on campus and through University of Michigan Health Research. There may be unique characteristics associated with men who take part in a study advertising for men

compared to men who were not explicitly targeted for recruitment (i.e., psychology subject pool). Being made aware of gender through recruitment materials may raise men's suspicions about the true intent of the research.

An additional limitation includes the lack of external validity from online samples and the university subject pool. Concern has been expressed that online recruitment produces sample bias because demographics of online survey participants may differ from the population as a whole (Hunter, 2012). Concern has also been raised regarding diversity of university subject pools compared to the general population (Paolacci, Chandler, & Ipeirotis, 2010). Researchers should interpret results from the present set of studies (which was done in a specific environment, at a specific cultural moment, and with certain types of people) with caution as they may not extend to other contexts.

There were also differences between samples regarding connectedness with the Twitter Community. In Study 1, participants rated Twitter as an important part of their daily routine. Community connectedness was markedly lower in Study 2: the majority of participants did not rate Twitter as an important part of their daily routine. Differences in Twitter connectedness could impact the strength of results. For example, men who hold misogynistic beliefs may seek out Twitter to support people who share their views. Active Twitter users could therefore hold different masculinity and gender beliefs compared to non-active Twitter users.

It is important to highlight that effect sizes from the present set of studies are relatively small. One possible explanation is small sample size: recent guidelines suggest that to produce large interaction effects, sample size should roughly be four times the number suggested by GPower analysis (Giner-Sorolla, 2018). Another possible is explanation is the importance of context: men may support or perpetrate online gender harassment for a variety of reasons, and

these reasons may differ by social location. It is perhaps not surprising to obtain small effects given that there are several factors that may shape men's endorsement and perpetration of online gender harassment. The development of a more comprehensive model that accounts for multiple predictors of online gender harassment should be examined in future research.

Future researchers should consider extending participant characteristics (e.g., race, sexual orientation, and age) to clarify the relationship between masculinity threat and online gender harassment. Because manhood is widely viewed as a precarious status *across* cultures and contexts, diverse men (e.g., race, age, sexual orientation) may be motivated to reassert masculinity under conditions of threat. Arguably, other social identities may not impact endorsement or perpetration of online gender harassment as much as threats to gender identity. Future work should assess if diverse social identities moderate the relationship between masculinity threat and likelihood to gender harass in Twitter and Facebook.

Researchers may also want to employ a control condition that involves a threat separate from gender. For instance, participants in Study 2 and Study 3 were randomly assigned to receive one of two types of false feedback ostensibly comparing their scores with those of other men who had completed the same personality inventory (masculinity affirmed, masculinity threatened). It is possible that the *experience of threat* impacts men's attitudes and behaviors, irrespective of gender. This leads to pressing questions: would threatening another social identity—such as race—produce similar results? Or does masculinity threat uniquely impact endorsement and perpetration of gender harassment in Twitter and Facebook? The comparison of the masculinity condition to the control condition would help clarify the role of gender threat in men's attitudes and behaviors.

Finally, it would also be a fruitful future direction to diversify the social identities of harassment targets. Research indicates that online harassment is not only about gender, but are also often racist, with women of color experiencing more extreme forms of abuse (Poland, 2016). It would be interesting to explore whether women of color are targets of more gender harassment compared to white women who posted the same content on Twitter and Facebook. A long-term goal of this research would be to evaluate more empirically informed policies and interventions aimed at decreasing racist and sexist harassment in social media —which would offer tangible solutions to preventing online harassment in the future.

Concluding Remarks

As online technology continues to become integrated in our lives, the question that we must consider is: What motivates gender harassment in a networked era? The present research illustrated that masculinity threat and anonymity impacted men's support and use of gender harassment in social media. These results suggest that instead of viewing gender harassment as an outcome of trolling, there is a "particular political purpose of male [hostile] behavior: the silencing of women who dare to speak in the online public sphere" (Megarry, 2014, p. 53). We must work to develop more empirically informed policies that consider masculinity, power, and technology.

Footnotes

[9] It's important to highlight that the interactive effect between masculinity norm conformity and toxic disinhibition did not replicate in Study 2. Differences in Twitter engagement may account for this discrepancy. Recruitment for Study 1 required that men identify as active Twitter users while recruitment for Study 2 did not require active engagement. There were also differences in key predictors, with men in Study 1 reporting significantly higher toxic disinhibition and conformity to masculine norms compared to men in Study 2. Variations between samples indicate that there may be unique characteristics associated with men who are active in Twitter. For example, men who hold misogynistic beliefs may seek out Twitter to support people who share their views. Future work should clarify the importance of Twitter engagement in men's support and use of gender harassment.

References

Adam, A. (2005). Internet Dating, Cyberstalking and Internet Pornography: Gender and the Gaze. In *Gender, Ethics and Information Technology* (pp. 102–127). Palgrave Macmillan, London. https://doi.org/10.1057/9780230000520_6

Angwin, J., & Grassegger, H. (2017, June 6). Facebook's Secret Censorship Rules Protect White Men… [text/html]. Retrieved December 15, 2017, from https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms

Banet-Weiser, S., & Miltner, K. M. (2016). #MasculinitySoFragile: culture, structure, and networked misogyny. *Feminist Media Studies*, *16*(1), 171–174. https://doi.org/10.1080/14680777.2016.1120490

Barak, A. (2005). Sexual Harassment on the Internet. *Social Science Computer Review*, *23*(1), 77–92. https://doi.org/10.1177/0894439304271540

Barlett, J., Norrie, R., Patel, N., Rumpel, R., & Wibberley. (2014). *Misogyny on Twitter*. Demos. Retrieved from https://www.demos.co.uk/project/misogyny-on-twitter/

Bartow, A. (2009). Internet Defamation as Profit Center: The Monetization of Online Harassment. *Harvard Journal of Law and Gender*, 383–429.

Berke, D. S., Reidy, D. E., Miller, J. D., & Zeichner, A. (2017). Take it like a man: Gender-threatened men's experience of gender role discrepancy, emotion activation, and pain tolerance. *Psychology of Men & Masculinity*, *18*(1), 62–69. https://doi.org/10.1037/men0000036

Bernstein, A. (2015). Abuse and Harassment Diminish Free Speech. *Pace Law Review*, *35*, 1–29.

Biebly, W. (2013). *Gender and Work: Challenging Conventional Wisdom* (pp. 1–7). Cambridge, Massachusettes: Harvard Business School.

Blackwell, L., Chen, T., Schoenebeck, S., & Lampe, C. (2018). When Online Harassment is Perceived as Justified. *Association for the Advancement of Artificial Intelligence*, 10.

Blackwell, L., Dimond, J., Schoenebeck, S., & Lampe, C. (2017). Classification and Its Consequences for Online Harassment: Design Insights from HeartMob. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 1–19. https://doi.org/10.1145/3134659

Bosson, J. K., & Vandello, J. A. (2011). Precarious Manhood and Its Links to Action and Aggression. *Current Directions in Psychological Science*, *20*(2), 82–86. https://doi.org/10.1177/0963721411402669

Bosson, J. K., Vandello, J. A., Burnaford, R. M., Weaver, J. R., & Wasti, S. A. (2009). Precarious Manhood and Displays of Physical Aggression. *Personality and Social Psychology Bulletin*, *35*(5), 623–634. https://doi.org/10.1177/0146167208331161

Brascombe, N. R., Ellemers, N., Spears, R., & Doosje, E. J. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & E. J. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 35–58). England: Blackwell.

Budgeon, S. (2011). The Contradictions of Successful Femininity: Third-Wave Feminism, Postfeminism and 'New' Femininities. In *New Femininities* (pp. 279–292). Palgrave Macmillan, London. Retrieved from https://link.springer.com/chapter/10.1057/9780230294523_19

Chemaly, S. (2014). There's no comparing male and female harassment online. *Time Magazine*. Retrieved from http://time.com/3305466/male-female-harassment- online/

Chen, A. S. (1999). Lives at the center of the periphery, Lives at the periperhy of center: Chinese American Masculinities and Bargaining with Hegemony. *Gender & Society*, *13*(5), 584–607. https://doi.org/10.1177/089124399013005002

Cheng, C. (1999). "We choose not the compensate:" The "merit" discourse in the selection process, and Asian and Asian American men and their masculinity. In C. Cheng (Ed.), *Masculinities in Organizations*. Thousand Oaks, California: SAGE.

Cheryan, S., Schwartz Cameron, J., Katagiri, Z., & Monin, B. (2015). Manning Up. *Social Psychology*, *46*(4), 218–227. https://doi.org/10.1027/1864-9335/a000239

Cho, D., & Kwon, K. H. (2015). The impacts of identity verification and disclosure of social cues on flaming in online user comments. *Computers in Human Behavior*, *51*, 363–372. https://doi.org/10.1016/j.chb.2015.04.046

Chu, A. (2014, May 27). Your Princess Is in Another Castle: Misogyny, Entitlement, and Nerds. *The Daily Beast*. Retrieved from https://www.thedailybeast.com/articles/2014/05/27/your-princess-is-in-another-castle-misogyny-entitlement-and-nerds

Cialdini, R. B. (2007). Descriptive Social Norms as Underappreciated Sources of Social Control. *Psychometrika; Williamsburg*, *72*(2), 263–268. http://dx.doi.org.proxy.lib.umich.edu/10.1007/s11336-006-1560-6

Citron, D. K. (2014). *Hate Crimes in Cyberspace*. Cambridge, Massachusettes: Harvard University Press.

Cohn, A. M., Seibert, L. A., & Zeichner, A. (2009). The role of restrictive emotionality, trait anger, and masculinity threat in men's perpetration of physical aggression. *Psychology of Men & Masculinity*, *10*(3), 218–224. https://doi.org/10.1037/a0015151

Conley, T. D., Ziegler, A., & Moors, A. C. (2013). Backlash From the Bedroom: Stigma Mediates Gender Differences in Acceptance of Casual Sex Offers. *Psychology of Women Quarterly*, *37*(3), 392–407. https://doi.org/10.1177/0361684312467169

Connell, R. W. (1985). Theorising Gender. *Sociology*, *19*(2), 260–272. https://doi.org/10.1177/0038038585019002008

Connell, R. W., & Connell, R. (2005). *Masculinities: Second Edition*. Berkeley, California: University of California Press.

Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic Masculinity: Rethinking the Concept. *Gender & Society*, *19*(6), 829–859. https://doi.org/10.1177/0891243205278639

Coston, B. M., & Kimmel, M. (2012). Seeing Privilege Where It Isn't: Marginalized Masculinities and the Intersectionality of Privilege. *Journal of Social Issues*, *68*(1), 97–111. https://doi.org/10.1111/j.1540-4560.2011.01738.x

Courtenay, W. H. (2000). Constructions of masculinity and their influence on men's well-being: a theory of gender and health. *Social Science & Medicine*, *50*(10), 1385–1401. https://doi.org/10.1016/S0277-9536(99)00390-1

Crocker, L. (2017, June 5). The Sick Facebook Rooms That Dashed 10 Kids' Harvard Dreams. *The Daily Beast*. Retrieved from https://www.thedailybeast.com/the-obscene-memes-that-dashed-these-harvard-students-dreams

Dahl, J., Vescio, T., & Weaver, K. (2015). How Threats to Masculinity Sequentially Cause Public Discomfort, Anger, and Ideological Dominance Over Women. *Social Psychology*, *46*(4), 242–254. https://doi.org/10.1027/1864-9335/a000248

Dillon, K. P., & Bushman, B. J. (2015). Unresponsive or un-noticed?: Cyberbystander intervention in an experimental cyberbullying context. *Computers in Human Behavior*, *45*(Supplement C), 144–150. https://doi.org/10.1016/j.chb.2014.12.009

Donaldson, M. (1993). What Is Hegemonic Masculinity? *Theory and Society*, *22*(5), 643–657.

Dubrovsky, V. J., Kiesler, S., & Sethna, B. N. (1991). The Equalization Phenomenon: Status Effects in Computer-Mediated and Face-to-Face Decision-Making Groups. *Human–Computer Interaction*, *6*(2), 119–146. https://doi.org/10.1207/s15327051hci0602_2

Duggan, M. (2014, October 22). Experiencing Online Harassment. Pew Research Center.

Duggan, M. (2017, July 11). Online Harassment 2017. Retrieved April 2, 2018, from http://www.pewinternet.org/2017/07/11/online-harassment-2017/

Eisler, R. M., & Skidmore, J. R. (1987). Masculine Gender Role Stress: Scale Development and Component Factors in the Appraisal of Stressful Situations. *Behavior Modification*, *11*(2), 123–136. https://doi.org/10.1177/01454455870112001

Ellison, N. B., Gray, R., Lampe, C., & Fiore, A. T. (2014). Social capital and resource requests on Facebook. *New Media & Society*, *16*(7), 1104–1121. https://doi.org/10.1177/1461444814543998

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Facebook Community Standards. (2017). Retrieved December 15, 2017, from https://www.facebook.com/communitystandards

Fahs, B., & Gohr, M. (2012). Superpatriarchy Meets Cyberfeminism: Facebook, Online Gaming, and the New Social Genocide. *MP: An Online Feminist Journal*, *3*(6), 1–40.

Faludi, S. (2009). *Backlash: The Undeclared War Against American Women*. New York, New York: Crown/Archetype.

Fenstermaker, S., & West, C. (2002). *Doing Gender, Doing Difference: Inequality, Power, and Institutional Change*. New York, New York: Routledge.

Ferguson, M., & Ford, T. (2008). Disparagement humor: A theoretical and empirical review of psychoanalytic, superiority, and social identity theories. *Humor*, *21*(3), 283–312. https://doi-org.proxy.lib.umich.edu/10.1515/HUMOR.2008.014

Fitzgerald, L. F., Drasgow, F., Hulin, C. L., Gelfand, M. J., & Magley, V. J. (1997). Antecedents and consequences of sexual harassment in organizations: A test of an integrated model. *Journal of Applied Psychology*, *82*(4), 578–589. https://doi.org/10.1037/0021-9010.82.4.578

Ford, T. E., Boxer, C. F., Armstrong, J., & Edel, J. R. (2008). More Than "Just a Joke": The Prejudice-Releasing Function of Sexist Humor. *Personality and Social Psychology Bulletin*, *34*(2), 159–170. https://doi.org/10.1177/0146167207310022

Ford, T. E., Woodzicka, J. A., Triplett, S. R., Kochersberger, A. O., & Holden, C. J. (2014). Not all groups are equal: Differential vulnerability of social groups to the prejudice-releasing effects of disparagement humor. *Group Processes & Intergroup Relations*, *17*(2), 178–199. https://doi.org/10.1177/1368430213502558

Fox, J., Cruz, C., & Lee, J. Y. (2015). Perpetuating online sexism offline: Anonymity, interactivity, and the effects of sexist hashtags on social media. *Computers in Human Behavior*, *52*(Supplement C), 436–442. https://doi.org/10.1016/j.chb.2015.06.024

Fox, J., & Tang, W. Y. (2014). Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior*, *33*(Supplement C), 314–320. https://doi.org/10.1016/j.chb.2013.07.014

Franks, M. A. (2012). Sexual Harassment 2.0 Special Feature: Cyberlaw. *Maryland Law Review*, *71*, 655–704.

Funk, M. E., & Coker, C. R. (2016). She's Hot, for a Politician: The Impact of Objectifying Commentary on Perceived Credibility of Female Candidates. *Communication Studies*, *67*(4), 455–473. https://doi.org/10.1080/10510974.2016.1196380

Gardiner, B., Mansfield, M., Anderson, I., Holder, J., Louter, D., & Ulmanu, M. (2016, April 12). The dark side of Guardian comments. *The Guardian*. Retrieved from http://www.theguardian.com/technology/2016/apr/12/the-dark-side-of-guardian-comments

Gautam, T. (2012, October 17). There's No Glass Ceiling On Twitter. Retrieved December 18, 2017, from https://www.forbes.com/sites/forbeswomanfiles/2012/10/17/theres-no-glass-ceiling-on-twitter/#31d33194522e

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception: Classic Edition* (1 edition). New York, London: Psychology Press.

Gill, R. (2016). Post-postfeminism?: new feminist visibilities in postfeminist times. *Feminist Media Studies*, *16*(4), 610–630. https://doi.org/10.1080/14680777.2016.1193293

Gilmore, D. D. (1990). *Manhood in the Making: Cultural Concepts of Masculinity*. Cambridge, Massachusettes: Yale University Press.

Giner-Sorolla, R. (2018, January 24). Powering Your Interaction. Retrieved September 13, 2018, from https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/

Glick, P., Gangl, C., Gibb, S., Klumpner, S., & Weinberg, E. (2007). Defensive Reactions to Masculinity Threat: More Negative Affect Toward Effeminate (but not Masculine) Gay Men. *Sex Roles*, *57*(1–2), 55–59. https://doi.org/10.1007/s11199-007-9195-3

Goff, P. A., Di Leone, B. A. L., & Kahn, K. B. (2012). Racism leads to pushups: How racial discrimination threatens subordinate men's masculinity. *Journal of Experimental Social Psychology*, *48*(5), 1111–1116. https://doi.org/10.1016/j.jesp.2012.03.015

Goffman, E. (1963). *Stigma: notes on the management of spoiled identity*. Englewood Cliffs, New Jersey: Prentice-Hall.

Gray, J. A., & Ford, T. E. (2013). The role of social context in the interpretation of sexist humor. *Humor*, *26*(2), 277–293. https://doi.org/10.1515/humor-2013-0017

Greenwood, D., & Isbell, L. M. (2002). Ambivalent Sexism and the Dumb Blonde: Men's and Women's Reactions to Sexist Jokes. *Psychology of Women Quarterly*, *26*(4), 341–350. https://doi.org/10.1111/1471-6402.t01-2-00073

Greenwood, S., Perrin, R. & Duggan, M. (2016, November 11). Facebook usage and engagement is on the rise, while adoption of other platforms holds steady. Retrieved December 15, 2017, from http://www.pewinternet.org/2016/11/11/social-media-update-2016/

Gross, T. (2017, November 1). Columnist Lindy West Sees "Straight Line" From Trolls Who Targeted Her To Trump. Retrieved December 15, 2017, from https://www.npr.org/2017/11/01/561350775/columnist-lindy-west-sees-straight-line-from-trolls-who-targeted-her-to-trump

Haraway, D. (1991). *Simians, Cyborgs, and Women*. New York: Routledge.

Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research. Language, Behaviour, Culture*, *6*(2), 215–242. https://doi.org/10.1515/jplr.2010.011

Henry, N., & Powell, A. (2015). Beyond the 'sext': Technology-facilitated sexual violence and harassment against adult women. *Australian & New Zealand Journal of Criminology*, *48*(1), 104–118. https://doi.org/10.1177/0004865814524218

Hess, A. (2014, January 6). The Next Civil Rights Issue: Why Women Aren't Welcome on the Internet. Retrieved December 14, 2017, from https://psmag.com/social-justice/women-arent-welcome-internet-72170

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Higgins, E. T., Klein, R., & Strauman, T. (1985). Self-Concept Discrepancy Theory: A Psychological Model for Distinguishing among Different Aspects of Depression and Anxiety. *Social Cognition*, *3*(1), 51–76. https://doi.org/10.1521/soco.1985.3.1.51

Hitlan, R. T., Pryor, J. B., Hesson-McInnis, M. S., & Olson, M. (2009). Antecedents of Gender Harassment: An Analysis of Person and Situation Factors. *Sex Roles*, *61*(11–12), 794–807. https://doi.org/10.1007/s11199-009-9689-2

Hodson, G., & MacInnis, C. C. (2016). Derogating humor as a delegitimization strategy in intergroup contexts. *Translational Issues in Psychological Science*, *2*(1), 63–74. https://doi.org/10.1037/tps0000052

Hudson, L. (2014, May 15). Curbing Online Abuse Isn't Impossible. Here's Where We Start. Retrieved April 3, 2018, from https://www.wired.com/2014/05/fighting-online-harassment/

Hunt, & Gonsalkorale, K. (2014). Who Cares What She Thinks, What Does He Say? Links between Masculinity, In-Group Bonding and Gender Harassment. *Sex Roles*, *70*(1–2), 14–27. https://doi.org/10.1007/s11199-013-0324-x

Hunter, L. (2012). Challenging the reported disadvantages of e-questionnaires and addressing methodological issues of online data collection. *Nurse Researcher*, *20*(1), 11–20.

Hutchby, I. (2001). Technologies, Texts and Affordances. *Sociology*, *35*(2), 441–456. https://doi.org/10.1017/S0038038501000219

Hutchens, M. J., Cicchirillo, V. J., & Hmielowski, J. D. (2015). How could you think that?!?!: Understanding intentions to engage in political flaming. *New Media & Society*, *17*(8), 1201–1219. https://doi.org/10.1177/1461444814522947

Jackson, L. M. (2017, November 16). Does Twitter Really Want to Solve Its Harassment Problem? *The New Republic*. Retrieved from https://newrepublic.com/article/145853/twitter-really-want-solve-harassment-problem

Jane, E. A. (2014a). 'Back to the kitchen, cunt': speaking the unspeakable about online misogyny. *Continuum*, *28*(4), 558–570. https://doi.org/10.1080/10304312.2014.924479

Jane, E. A. (2014b). "Your a Ugly, Whorish, Slut:" Understanding e-bile. *Feminist Media Studies*, *14*(4), 531–546. https://doi.org/10.1080/14680777.2012.741073

Jane, E. A. (2016). *Misogyny Online: A Short (and Brutish) History*. Thousand Oaks, California: SAGE.

Jhaver, S., Ghoshal, S., Bruckman, A., & Gilbert, E. (2018). Online Harassment and Content Moderation: The Case of Blocklists. *ACM Trans. Comput.-Hum. Interact.*, *25*(2), 11–33.

Johnson, L. (2016, June 20). Why Big Brands Are Suddenly Getting Cozy With Reddit. Retrieved April 2, 2018, from http://www.adweek.com/digital/why-big-brands-are-suddenly-getting-cozy-reddit-172053/

Junco, R. (2012). Too much face and not enough books: The relationship between multiple indices of Facebook use and academic performance. *Computers in Human Behavior*, *28*(1), 187–198. https://doi.org/10.1016/j.chb.2011.08.026

Katz, J. (2006). *The Macho Paradox: Why Some Men Hurt Women and and How All Men Can Help* (3.2.2006 edition). Naperville, Ill: Sourcebooks.

Kehily, M. J., & Nayak, A. (1997). "Lads and Laughter": Humour and the production of heterosexual hierarchies. *Gender and Education*, *9*(1), 69–88. https://doi.org/10.1080/09540259721466

Khan, K., Spencer, K., & Glaser, J. (2013). Online Prejudice and Discrimination: From Dating to Hating - Oxford Scholarship. In Y. Amichai-Hamburger (Ed.), *The Social Net: Understanding our online behavior*. Oxford Scholarship Online. Retrieved from http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199639540.001.0001/acprof-9780199639540-chapter-011

Kimmel, M. (1994). Foreword. In *Theorizing Masculinities* (pp. vii–ix). London: Sage Publications Ltd.

Kimmel, M. (2008). *Guyland: The Perilous World Where Boys Become Men*. New York, New York: HaperCollins.

Kimmel, M. (2013, July 16). Opinion | Fired for Being Beautiful. *The New York Times*. Retrieved from https://www.nytimes.com/2013/07/17/opinion/fired-for-being-beautiful.html

Klonick, K. (2015). *Re-Shaming the Debate: Social Norms, Shame, and Regulation in an Internet Age* (SSRN Scholarly Paper No. ID 2638693). Rochester, NY: Social Science Research Network. Retrieved from https://papers-ssrn-com.proxy.lib.umich.edu/abstract=2638693

Klonick, K. (2016, August 24). You'll Never Guess This One Crazy Thing Governs Online Speech. *Slate*. Retrieved from http://www.slate.com/articles/technology/future_tense/2016/08/free_speech_is_the_wrong_way_to_think_about_twitter_and_facebook.html

Knight-McCord, J., Cleary, D., Grant, N., Herron, A., Jumbo, S., Tiffany, L., … Smith, R. (2016). What social media sites do college students use most? *Journal of Undergraduate Ethnic Minority Psychology*, *2*, 21–26.

Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, *28*(2), 434–443. https://doi.org/10.1016/j.chb.2011.10.014

Lea, M., & Spears, R. (1991). Computer-mediated communication, de-individuation and group decision-making. *International Journal of Man-Machine Studies*, *34*(2), 283–301. https://doi.org/10.1016/0020-7373(91)90045-9

Lenhart, A., Ybarra, M., & Price-Feeney, M. (2016). *Online Harassment, Digital Abuse, and Cyberstalking in America*. Data and Society Research Institute.

Leskinen, E. A., & Cortina, L. M. (2014). Dimensions of Disrespect: Mapping and Measuring Gender Harassment in Organizations. *Psychology of Women Quarterly*, *38*(1), 107–123. https://doi.org/10.1177/0361684313496549

Leskinen, E. A., Rabelo, V. C., & Cortina, L. M. (2015). Gender stereotyping and harassment: A "catch-22" for women in the workplace. *Psychology, Public Policy, and Law*, *21*(2), 192–204. https://doi.org/10.1037/law0000040

Levant, R. F., Rankin, T. J., Williams, C. M., Hasan, N. T., & Smalley, K. B. (2010). Evaluation of the factor structure and construct validity of scores on the Male Role Norms Inventory—Revised (MRNI-R). *Psychology of Men & Masculinity*, *11*(1), 25–37. https://doi.org/10.1037/a0017637

Lewis, R., Rowe, M., & Wiper, C. (2017). Online Abuse of Feminists as An Emerging form of Violence Against Women and Girls. *The British Journal of Criminology*, *57*(6), 1462–1481. https://doi.org/10.1093/bjc/azw073

Lumsden, K., & Morgan, H. (2017). Media framing of trolling and online abuse: silencing strategies, symbolic violence, and victim blaming. *Feminist Media Studies*, *17*(6), 926–940. https://doi.org/10.1080/14680777.2017.1316755

Maass, A., Cadinu, M., Guarnieri, G., & Grasselli, A. (2003). Sexual harassment under social identity threat: The computer harassment paradigm. *Journal of Personality and Social Psychology*, *85*(5), 853–870. https://doi.org/10.1037/0022-3514.85.5.853

Mantilla, K. (2013). Gendertrolling: Misogyny Adapts to New Media. *Feminist Studies*, *39*(2), 563–570.

Mantilla, K. (2015). *Gendertrolling: How Misogyny Went Viral: How Misogyny Went Viral*. Santa Barbara, California: ABC-CLIO.

Mao, J. (2014). Social media for learning: A mixed methods study on high school students' technology affordances and perspectives. *Computers in Human Behavior*, *33*(Supplement C), 213–223. https://doi.org/10.1016/j.chb.2014.01.002

Marcotte, A. (2012, June 13). Online Misogyny: Can't Ignore It, Can't Not Ignore It. *Slate*. Retrieved from http://www.slate.com/blogs/xx_factor/2012/06/13/online_misogyny_reflects_women_s_realities_though_in_a_cruder_way_than_is_customary_offline_.html

Marwick, A. E., & Miller, R. (2014). *Online Harassment, Defamation, and Hateful Speech: A Primer of the Legal Landscape* (Fordham Center on Law and Information Policy Report) (pp. 1–75). New York: Fordham Center on Law and Information.

Megarry, J. (2014). Online incivility or sexual harassment? Conceptualising women's experiences in the digital age. *Women's Studies International Forum*, *47*(Part A), 46–55. https://doi.org/10.1016/j.wsif.2014.07.012

Meyer, R., & Cukier, M. (2006). Assessing the Attack Threat due to IRC Channels. In *International Conference on Dependable Systems and Networks (DSN'06)* (pp. 467–472). https://doi.org/10.1109/DSN.2006.12

Moor, P. J., Heuvelman, A., & Verleur, R. (2010). Flaming on YouTube. *Computers in Human Behavior*, *26*(6), 1536–1546. https://doi.org/10.1016/j.chb.2010.05.023

Moore, T., & Stuart, G. (2005). A Review of the Literature on Masculinity and Partner Violence. *Psychology of Men & Masculinity, 6*(1), 46–61. http://dx.doi.org.proxy.lib.umich.edu/10.1037/1524-9220.6.1.46

Moreton-Robinson, A. M. (2000). Duggaibah, or 'place of whiteness' : Australian feminist and race. In J. Docker & G. Fischer (Eds.), *Race, Colour and Identity in Australia and New Zealand* (pp. 240–255). Sydney: University of New South Wales Press.

Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: Status incongruity and backlash against modest men. *Psychology of Men & Masculinity*, *11*(2), 140–151. https://doi.org/10.1037/a0018093

Munsch, C. L., & Willer, R. (2012). The Role of Gender Identity Threat in Perceptions of Date Rape and Sexual Coercion. *Violence Against Women*, *18*(10), 1125–1146. https://doi.org/10.1177/1077801212465151

Nathanson, H. (2017, June 5). Harvard Rescinds Acceptances for At Least Ten Students for Obscene Memes | News | The Harvard Crimson. Retrieved December 15, 2017, from http://www.thecrimson.com/article/2017/6/5/2021-offers-rescinded-memes/

O'Connor, E. C., Ford, T. E., & Banos, N. C. (2017). Restoring Threatened Masculinity: The Appeal of Sexist and Anti-Gay Humor. *Sex Roles*, *77*(9–10), 567–580. https://doi.org/10.1007/s11199-017-0761-z

O'Malley, H. (2015, April 10). 3 ways to reform toxic masculinity after Gamergate. Retrieved May 5, 2017, from https://www.dailydot.com/via/reforming-toxic-masculinity-gamergate/

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*, 411–419.

Parent, M. C., & Moradi, B. (2011). An abbreviated tool for assessing conformity to masculine norms: Psychometric properties of the Conformity to Masculine Norms Inventory-46. *Psychology of Men & Masculinity*, *12*(4), 339–353. https://doi.org/10.1037/a0021904

Parmelee, J. H., & Bichard, S. L. (2012). *Politics and the Twitter Revolution: How Tweets Influence the Relationship Between Political Leaders and the Public*. Lexington Books.

Pascoe, C. J. (2011). *Dude, You're a Fag: Masculinity and Sexuality in High School*. Berkeley, California: University of California Press.

Pascoe, C. J., & Bridges, T. (Eds.). (2016). *Exploring Masculinities : Identity, Inequality, Continuity, and Change*. New York: Oxford University Press.

Pater, J. A., Kim, M. K., Mynatt, E. D., & Fiesler, C. (2016). Characterizations of Online Harassment: Comparing Policies Across Social Media Platforms (pp. 369–374). ACM Press. https://doi.org/10.1145/2957276.2957297

Phillips, W. (2015). *This Is Why We Can't Have Nice Things: Mapping the Relationship Between Online Trolling and Mainstream Culture*. Cambridge, Massachusetts: MIT Press.

Pleck, J. H. (1977). The psychology of sex roles: Traditional and new views. In A. Scott & L. Carter (Eds.), *Women and Men: Changing, Roles, Attitudes and Perceptions* (pp. 181–201). New York: Praeger.

Pleck, J. H. (1981). *The myth of masculinity.* Cambridge, Massachusetts: MIT Press.

Pleck, J. H. (1995). The gender role strain paradigm: An update. In R. H. Levant & W. S. Pollack (Eds.), *A new psychology of men* (pp. 11–32). New York: Basic Books.

Poland, B. (2016). *Haters: Harassment, Abuse, and Violence Online*. Lincoln, Nebraska: University of Nebraska Press.

Postmes, T., & Spears, R. (2002). Behavior Online: Does Anonymous Computer Communication Reduce Gender Inequality? *Personality and Social Psychology Bulletin*, *28*(8), 1073–1083.

Postmes, T., Spears, R., & Lea, M. (1998). Breaching or Building Social Boundaries? SIDE-Effects of Computer-Mediated Communication. *Communication Research*, *25*(6), 689–715. https://doi-org.proxy.lib.umich.edu/10.1177/009365098025006006

Postmes, T., Spears, R., & Lea, M. (1999). Social identity, group norms, and "deindividuation": Lessons from computer-mediated communication for social influence in the group. In R. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content*. Oxford: Blackwell.

Postmes, T., Spears, R., Sakhel, K., & de Groot, D. (2001). Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior. *Personality and Social Psychology Bulletin*, *27*(10), 1243–1254. https://doi.org/10.1177/01461672012710001

Rainie, L., Anderson, J., & Albright, J. (2017, March 29). The Future of Free Speech, Trolls, Anonymity and Fake News Online. Retrieved April 3, 2018, from http://www.pewinternet.org/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/

Reicher, S. D., Spears, R., & Postmes, T. (1995). A Social Identity Model of Deindividuation Phenomena. *European Review of Social Psychology*, *6*(1), 161–198. https://doi.org/10.1080/14792779443000049

Reidy, D. E., Brookmeyer, K. A., Gentile, B., Berke, D. S., & Zeichner, A. (2016). Gender Role Discrepancy Stress, High-Risk Sexual Behavior, and Sexually Transmitted Disease. *Archives of Sexual Behavior*, *45*(2), 459–465. https://doi.org/10.1007/s10508-014-0413-0

Reidy, D. E., Smith-Darden, J. P., Cortina, K. S., Kernsmith, R. M., & Kernsmith, P. D. (2015). Masculine Discrepancy Stress, Teen Dating Violence, and Sexual Violence Perpetration Among Adolescent Boys. *Journal of Adolescent Health*, *56*(6), 619–624. https://doi.org/10.1016/j.jadohealth.2015.02.009

Ritter, B. A. (2014). Deviant Behavior in Computer-Mediated Communication: Development and Validation of a Measure of Cybersexual Harassment. *Journal of Computer-Mediated Communication*, *19*(2), 197–214. https://doi.org/10.1111/jcc4.12039

Rowe, I. (2015). Deliberation 2.0: Comparing the Deliberative Quality of Online News User Comments Across Platforms. *Journal of Broadcasting & Electronic Media*, *59*(4), 539–555. https://doi.org/10.1080/08838151.2015.1093482

Rudman, L. A., & Fairchild, K. (2004). Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance. *Journal of Personality and Social Psychology*, *87*(2), 157–176. https://doi.org/10.1037/0022-3514.87.2.157

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Salter, M. (2017). From geek masculinity to Gamergate: the technological rationality of online abuse. *Crime, Media, Culture*. Advance online publication. https://doi.org/10.1177/1741659017690893

Sandberg, S. (2013). *Lean in: Women, work, and the will to lead*. New York: Random House.

Sarkeesian, A. (2012, June 7). Harassment, Misogyny and Silencing on YouTube. Retrieved December 15, 2017, from https://feministfrequency.com/2012/06/07/harassment-misogyny-and-silencing-on-youtube/

Schmitt, M. T., & Branscombe, N. R. (2001). The Good, the Bad, and the Manly: Threats to One's Prototypicality and Evaluations of Fellow In-Group Members. *Journal of Experimental Social Psychology*, *37*(6), 510–517. https://doi.org/10.1006/jesp.2001.1476

Scott, Allison, Kapor, F., & Onovakpuri, U. (2017). *Tech Leavers Study* (p. 27). Oakland, CA: Kapor Center for Social Impact.

Shifman, L. (2012). An anatomy of a YouTube meme. *New Media & Society*, *14*(2), 187–203. https://doi.org/10.1177/1461444811412160

Shifman, L., & Lemish, D. (2010). Between Feminism and Fun(ny)mism. *Information, Communication & Society*, *13*(6), 870–891. https://doi.org/10.1080/13691180903490560

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*, *7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Sussman, N. M., & Tyson, D. H. (2000). Sex and power: gender differences in computer-mediated interactions. *Computers in Human Behavior*, *16*(4), 381–394. https://doi.org/10.1016/S0747-5632(00)00020-0

Tajfel, H., & Turner, J. C. (1986). The social identity theory of inter group behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations*. Chicago: Nelson.

Thomae, M., & Pina, A. (2015). Sexist humor and social identity: the role of sexist humor in men's in-group cohesion, sexual harassment, rape proclivity, and victim blame. *HUMOR*, *28*(2), 187–204. https://doi.org/10.1515/humor-2015-0023

Thompson, E. H., & Bennett, K. M. (2015). Measurement of masculinity ideologies: A (critical) review. *Psychology of Men & Masculinity*, *16*(2), 115–133. https://doi.org/10.1037/a0038609

Thompson, E. H., Pleck, J. H., & Ferrera, D. L. (1992). Men and masculinities: Scales for masculinity ideology and masculinity-related constructs. *Sex Roles*, *27*(11–12), 573–607. https://doi.org/10.1007/BF02651094

Udris, R. (2014). Cyberbullying among high school students in Japan: Development and validation of the Online Disinhibition Scale. *Computers in Human Behavior*, *41*(Supplement C), 253–261. https://doi.org/10.1016/j.chb.2014.09.036

Valenti, J. (2014). *Full Frontal Feminism: A Young Woman's Guide to Why Feminism Matters*. Da Capo Press.

Valenti, J. (2015, August 29). Anita Sarkeesian interview: "The word 'troll' feels too childish. This is abuse." *The Guardian*. Retrieved from http://www.theguardian.com/technology/2015/aug/29/anita-sarkeesian-gamergate-interview-jessica-valenti

Vandello, J. A., & Bosson, J. K. (2013). Hard won and easily lost: A review and synthesis of theory and research on precarious manhood. *Psychology of Men & Masculinity*, *14*(2), 101–113. https://doi.org/10.1037/a0029826

Vandello, J. A., Bosson, J. K., Cohen, D., Burnaford, R. M., & Weaver, J. R. (2008). Precarious manhood. *Journal of Personality and Social Psychology*, *95*(6), 1325–1339. https://doi.org/10.1037/a0012453

Vikery, J. R., Everback, T., Blackwell, L., Franks, M. A., Friedman, B., Gibbons, S., … Massanari, A. (2018). Conclusion: What Can We Do About Mediated Misogyny? In J. R. Vickery & T. Everback (Eds.), *Mediating Misogyny: Gender, Technology, and Harassment* (pp. 399–422). London: Palgrave Macmillan.

Weaver, K. S., & Vescio, T. K. (2015). The Justification of Social Inequality in Response to Masculinity Threats. *Sex Roles*, *72*(11–12), 521–535. https://doi.org/10.1007/s11199-015-0484-y

West, C., & Zimmerman, D. H. (1987). Doing Gender. *Gender & Society*, *1*(2), 125–151. https://doi.org/10.1177/0891243287001002002

Willer, R., Rogalin, C. L., Conlon, B., & Wojnowicz, M. T. (2013). Overdoing Gender: A Test of the Masculine Overcompensation Thesis. *American Journal of Sociology*, *118*(4), 980–1022. https://doi.org/10.1086/668417

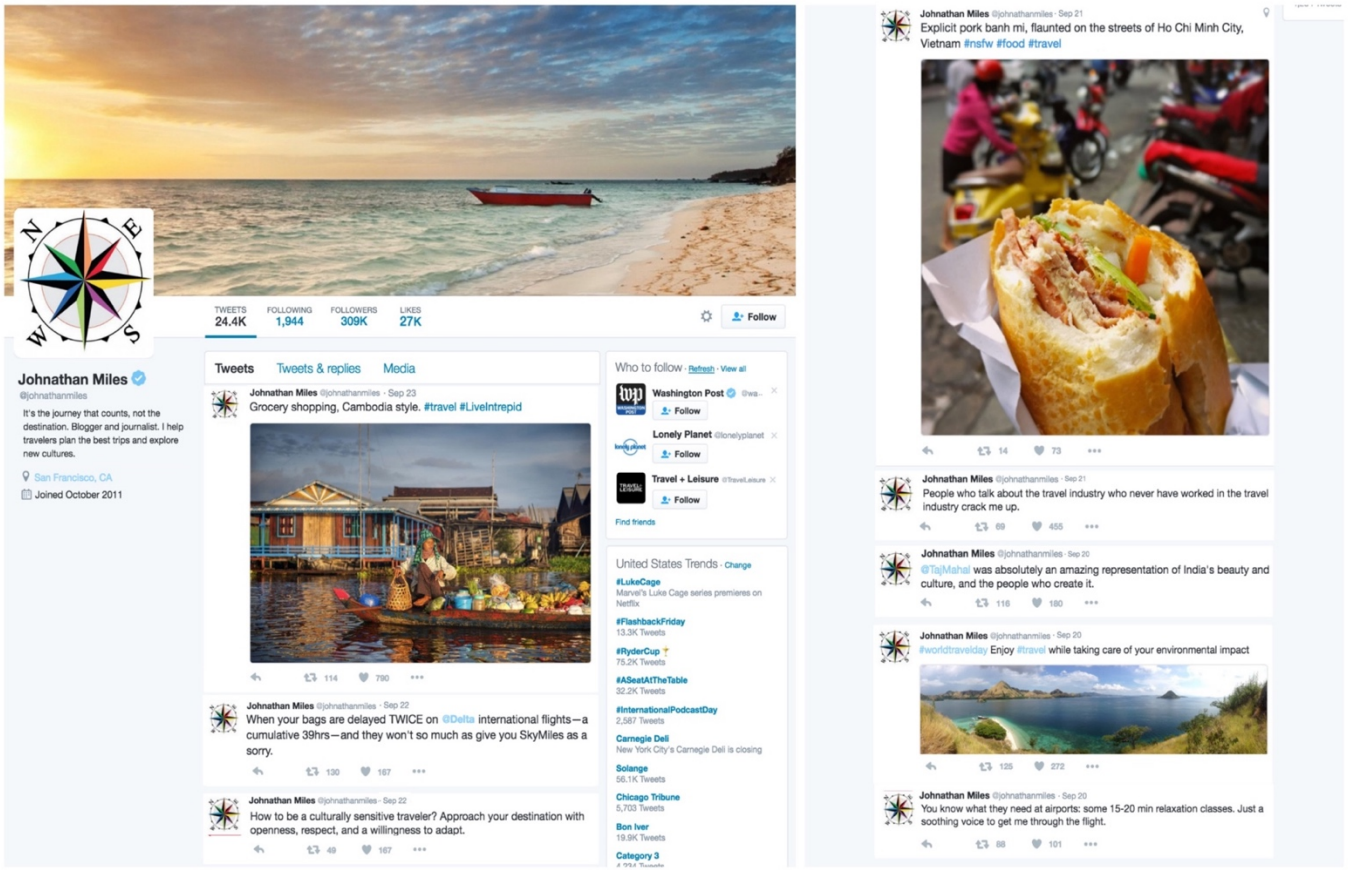
Wohn, D. Y., Fiesler, C., Hemphill, L., De Choudhury, M., & Matias, J. N. (2017). How to Handle Online Risks?: Discussing Content Curation and Moderation in Social Media (pp. 1271–1276). ACM Press. https://doi.org/10.1145/3027063.3051141

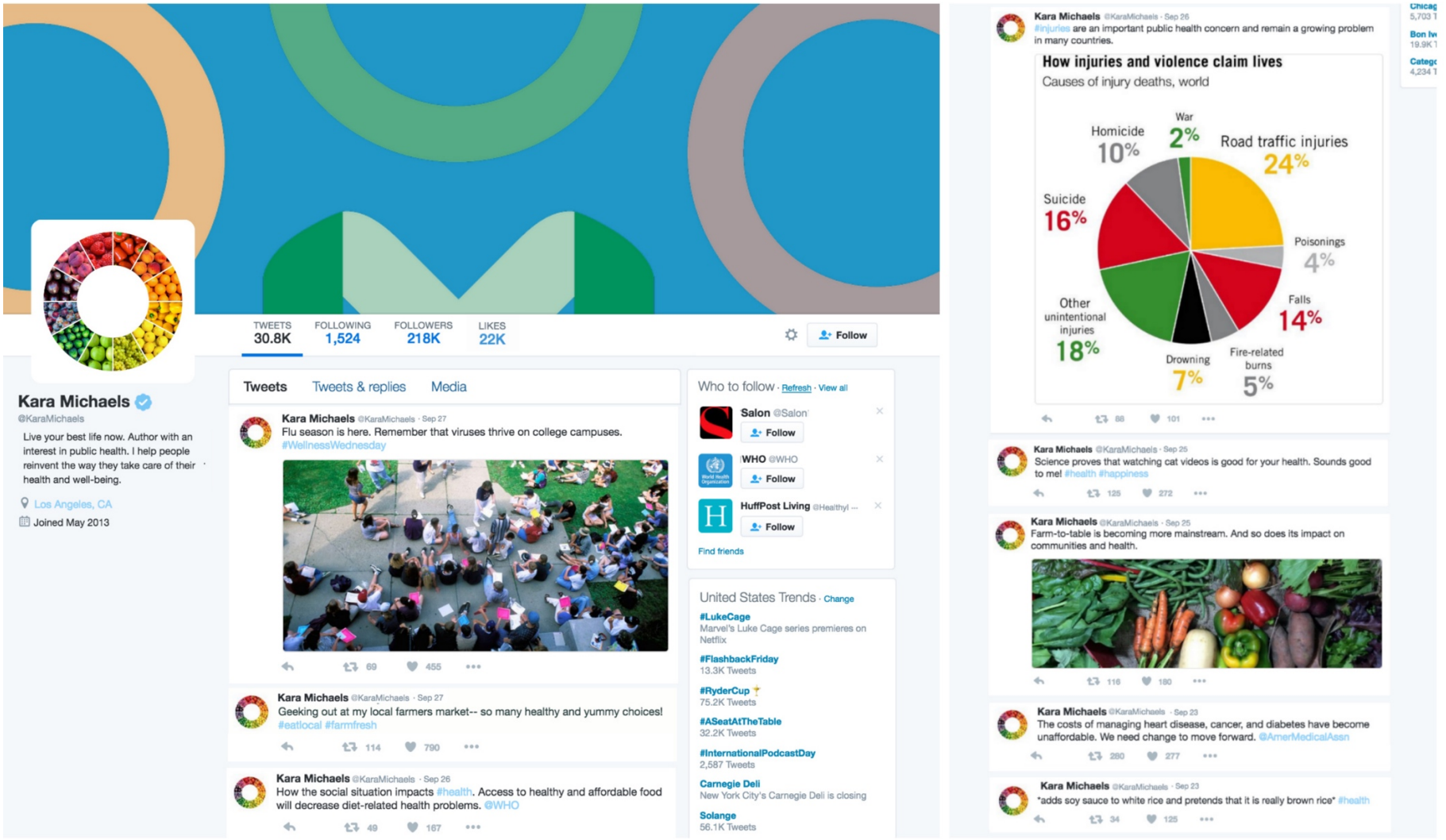
Woodzicka, J. A., & LaFrance, M. (2001). Real Versus Imagined Gender Harassment. *Journal of Social Issues*, *57*(1), 15–30. https://doi.org/10.1111/0022-4537.00199

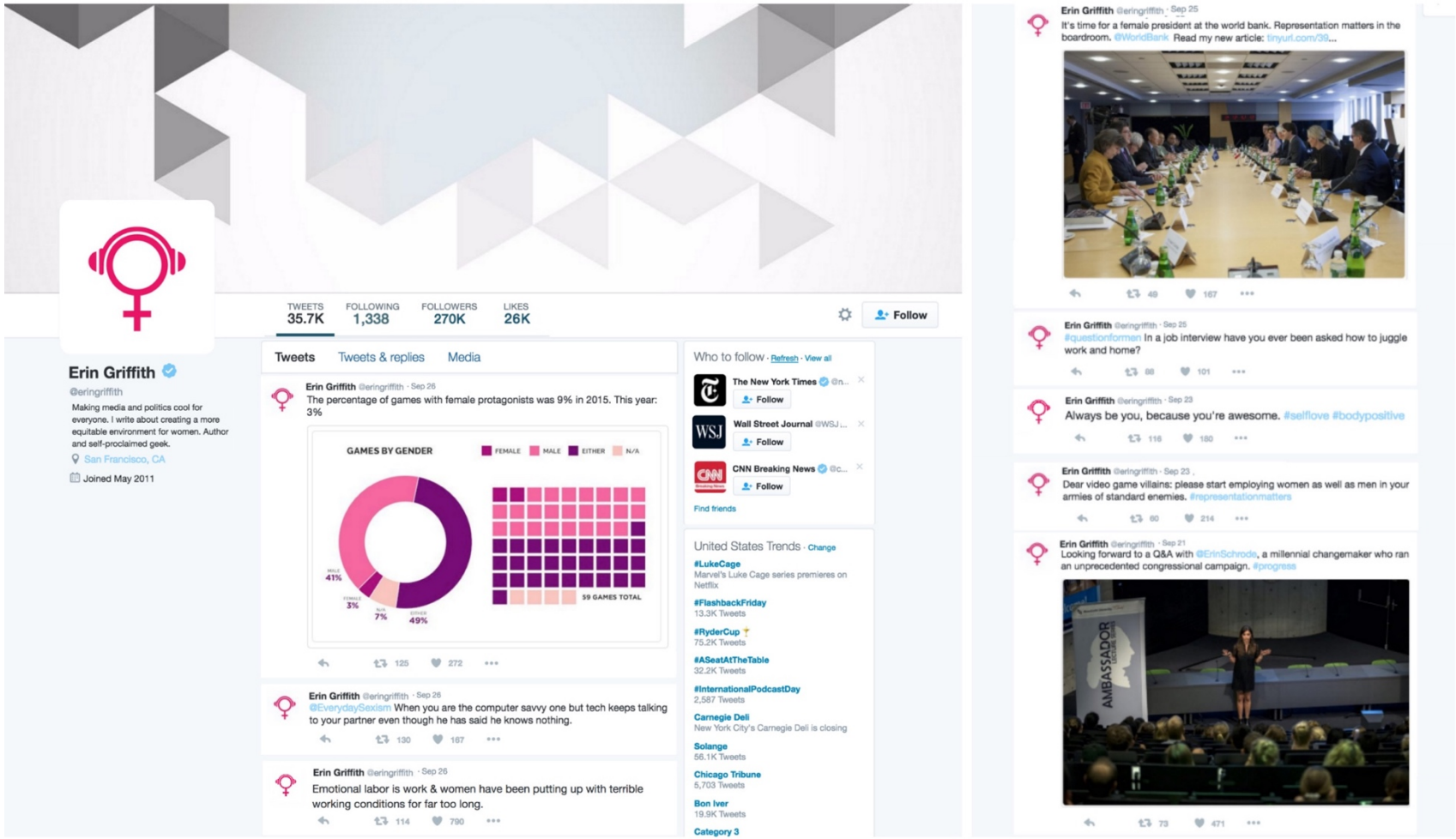
Appendices

Appendix A

Figure 12. Twitter profiles used in Study 1 and Study 2




TWEETS 30.8K
FOLLOWING 1,524
FOLLOWERS 218K
LIKES 22K
Follow
Kara Michaels
@KaraMichaels
Live your best life now. Author with an interest in public health. I help people reinvent the way they take care of their health and well-being.
Los Angeles, CA
Joined May 2013
Tweets
Tweets & replies
Media
Kara Michaels @KaraMichaels · Sep 27
Flu season is here. Remember that viruses thrive on college campuses. #WellnessWednesday
69 455
Kara Michaels @KaraMichaels · Sep 27
Geeking out at my local farmers market-- so many healthy and yummy choices! #eatlocal #farmfresh
114 790
Kara Michaels @KaraMichaels · Sep 26
How the social situation impacts #health. Access to healthy and affordable food will decrease diet-related health problems. @WHO
49 167
Who to follow · Refresh · View all
Salon @Salon
Follow
WHO @WHO
Follow
HuffPost Living @Healthyl ...
Follow
Find friends
United States Trends · Change
#LukeCage
Marvel's Luke Cage series premieres on Netflix
#FlashbackFriday
13.3K Tweets
#RyderCup
75.2K Tweets
#ASeatAtTheTable
32.2K Tweets
#InternationalPodcastDay
2,587 Tweets
Carnegie Deli
New York City's Carnegie Deli is closing
Solange
56.1K Tweets
Kara Michaels @KaraMichaels · Sep 26
#injuries are an important public health concern and remain a growing problem in many countries.
How injuries and violence claim lives
Causes of injury deaths, world
Homicide 10%
War 2%
Road traffic injuries 24%
Suicide 16%
Poisonings 4%
Other unintentional injuries 18%
Falls 14%
Drowning 7%
Fire-related burns 5%
88 101
Kara Michaels @KaraMichaels · Sep 25
Science proves that watching cat videos is good for your health. Sounds good to me! #health #happiness
125 272
Kara Michaels @KaraMichaels · Sep 25
Farm-to-table is becoming more mainstream. And so does its impact on communities and health.
116 180
Kara Michaels @KaraMichaels · Sep 23
The costs of managing heart disease, cancer, and diabetes have become unaffordable. We need change to move forward. @AmerMedicalAssn
280 277
Kara Michaels @KaraMichaels · Sep 23
"adds soy sauce to white rice and pretends that it is really brown rice" #health
34 125


TWEETS 35.7K
FOLLOWING 1,338
FOLLOWERS 270K
LIKES 26K
Follow
Erin Griffith
@eringriffith
Making media and politics cool for everyone. I write about creating a more equitable environment for women. Author and self-proclaimed geek.
San Francisco, CA
Joined May 2011
Tweets
Tweets & replies
Media
Erin Griffith @eringriffith · Sep 26
The percentage of games with female protagonists was 9% in 2015. This year: 3%
GAMES BY GENDER
FEMALE
MALE
EITHER
N/A
41%
3%
7%
49%
59 GAMES TOTAL
125
272
Erin Griffith @eringriffith · Sep 26
@EverydaySexism When you are the computer savvy one but tech keeps talking to your partner even though he has said he knows nothing.
130
167
Erin Griffith @eringriffith · Sep 26
Emotional labor is work & women have been putting up with terrible working conditions for far too long.
114
790
Who to follow · Refresh · View all
The New York Times @n..
Follow
Wall Street Journal @WSJ...
Follow
CNN Breaking News @c...
Follow
Find friends
United States Trends · Change
#LukeCage
Marvel's Luke Cage series premieres on Netflix
#FlashbackFriday
13.3K Tweets
#RyderCup
75.2K Tweets
#ASeatAtTheTable
32.2K Tweets
#InternationalPodcastDay
2,587 Tweets
Carnegie Deli
New York City's Carnegie Deli is closing
Solange
56.1K Tweets
Chicago Tribune
5,703 Tweets
Bon Iver
19.9K Tweets
Category 3
Erin Griffith @eringriffith · Sep 25
It's time for a female president at the world bank. Representation matters in the boardroom. @WorldBank Read my new article: tinyurl.com/39...
49
167
Erin Griffith @eringriffith · Sep 25
#questionformen In a job interview have you ever been asked how to juggle work and home?
88
101
Erin Griffith @eringriffith · Sep 23
Always be you, because you're awesome. #selflove #bodypositive
116
180
Erin Griffith @eringriffith · Sep 23
Dear video game villains: please start employing women as well as men in your armies of standard enemies. #representationmatters
60
214
Erin Griffith @eringriffith · Sep 21
Looking forward to a Q&A with @ErinSchrode, a millennial changemaker who ran an unprecedented congressional campaign. #progress
AMBASSADOR
73
471

Appendix B

Figure 13. Tweets (supportive and inflammatory tweets) used in Study 1 and Study 2

Note. Response items from the tweets 2-4 were combined to create an endorsement of harassment score.

Appendix C

Study 1 Measures

Gender Role Discrepancy and Discrepancy Stress Items

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

1. I am less masculine than the average guy. (GRD)
2. Compared to my guy friends, I am not very masculine. (GRD)
3. I wish I was more ‘‘manly.’’ (GRDS)
4. Most women I know would say that I am not as masculine as my friends. (GRD)
5. I wish I was interested in things that other guys find interesting. (GRDS)
6. Most women would consider me to be less masculine than the typical guy. (GRD)
7. I worry that people judge me because I am not like the typical man. (GRDS)
8. Most guys would think I am not very masculine compared to them. (GRD)
9. Sometimes I worry about my masculinity. (GRDS)
10. I worry that women find me less attractive because I’m not as macho as other guys. (GRDS)

Online Disinhibition Scale

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

Benign disinhibition

1. It is easier to connect with others through ICTs (information and communication technologies) than talking in person.
2. The Internet is anonymous so it is easier for me to express my true feelings or thoughts.
3. It is easier to write things online that would be hard to say in real life because you don’t see the other’s face.
4. It is easier to communicate online because you can reply anytime you like.
5. I have an image of the other person in my head when I read their e-mail or messages online.
6. I feel like a different person online.
7. I feel that online I can communicate on the same level with others who are older or have higher status.

Toxic disinhibition
8. I don't mind writing insulting things about others online, because it's anonymous.
9. It is easy to write insulting things online because there are no repercussions.
10. There are no rules online therefore you can do whatever you want.
11. Writing insulting things online is not bullying.

Conformity to Masculine Norms (11-item shortened form)

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

1. I tend to keep my feelings to myself.
2. In general, I will do anything to win.
3. If I could, I would frequently change sexual partners.
4. Sometimes violent action is necessary.
5. I hate asking for help.
6. I frequently put myself in risky situations.
7. In general, I control the women in my life.
8. In general, I must get my way.
9. My work is the most important part of my life.
10. It feels good to be important.
11. I would be furious if someone thought I was gay.

Twitter Intensity Scale

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

1. Twitter is part of my everyday activity.
2. I am proud to tell people that I am on Twitter.
3. Twitter has become part of my daily routine.
4. I feel out of touch when I haven't logged onto Twitter in awhile.
5. I feel I am part of the Twitter community.
6. I would be sorry if Twitter shut down.

Twitter Media Use

0-100 minutes, sliding scale

1. In the past week, on average, approximately how many MINUTES PER DAY have you spent actively using Twitter?

2. In the past week, on average, approximately how many MINUTES PER DAY have you spent:

- Posting Original Tweets
- Reading Tweets
- Responding to Tweets
- Retweeting

Appendix D

Study 2 Measures

Disappointment with Results

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

1. I am disappointed with my results.
2. I feel that my results accurately reflect me as a person.

Public Shame

Imagine that we publish your personality scores in social media. When you think about your name and scores being published in social media, how _____ do you feel?

Unconcerned . . . . . Concerned
Not Nervous . . . . . Nervous
Unapologetic . . . . . Defensive
Happy . . . . . Depressed
Calm . . . . . Upset
Confident . . . . . Insecure
Joyful . . . . . Sad
Collected . . . . . Angry

Conformity to Masculine Norms (46-item shortened form)

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

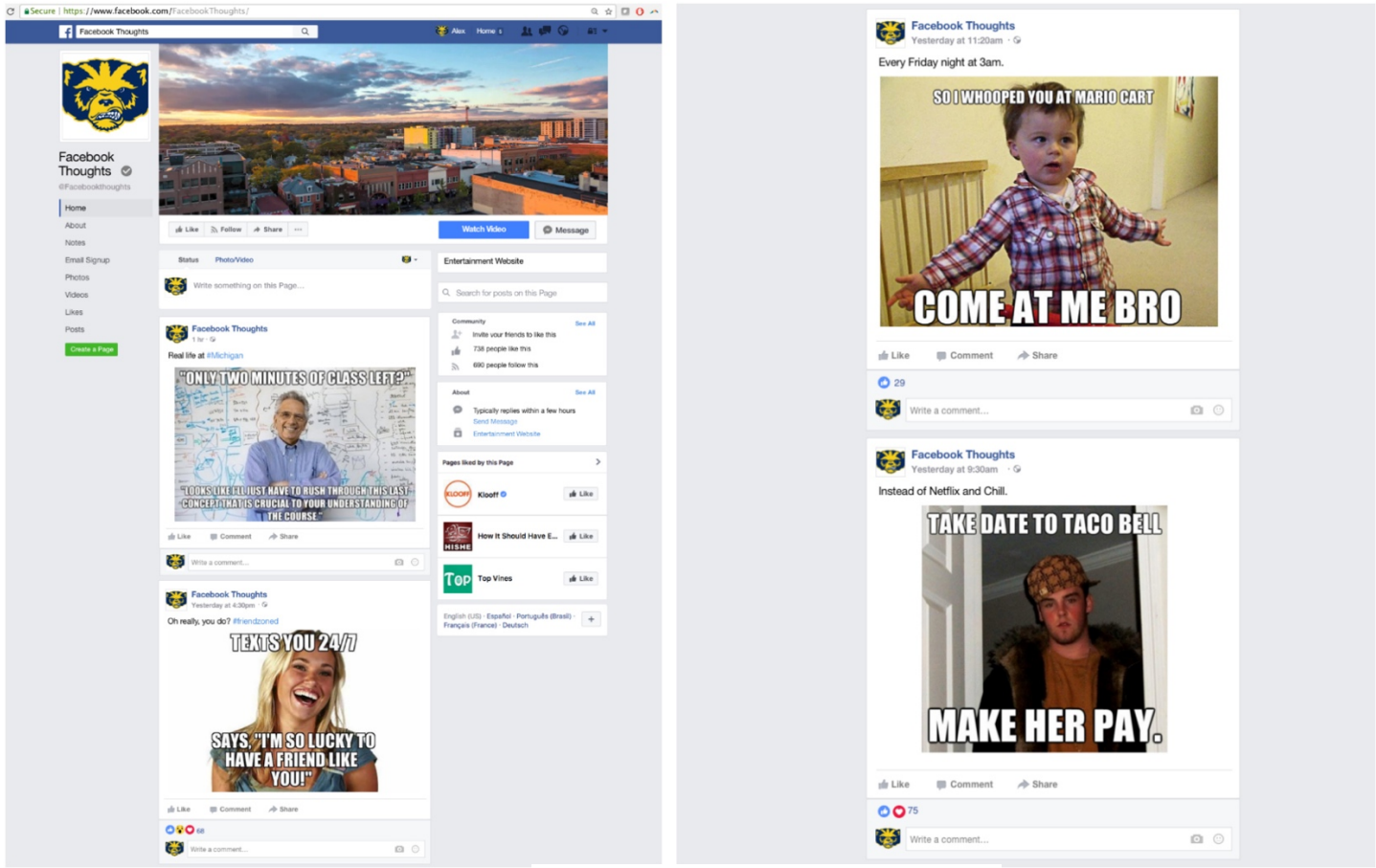
Instructions: Thinking about your own actions, feelings and beliefs, please indicate how much you personally agree or disagree with each statement. There are no right or wrong responses to the statements. You should give the responses that most accurately describe your personal actions, feelings and beliefs. It is best if you respond with your first impression when answering.

1. In general, I will do anything to win.
2. If I could, I would frequently change sexual partners.
3. I hate asking for help.
4. I believe that violence is never justified. (reverse)
5. Being thought of as gay is not a bad thing. (reverse)
6. In general, I do not like risky situations. (reverse)

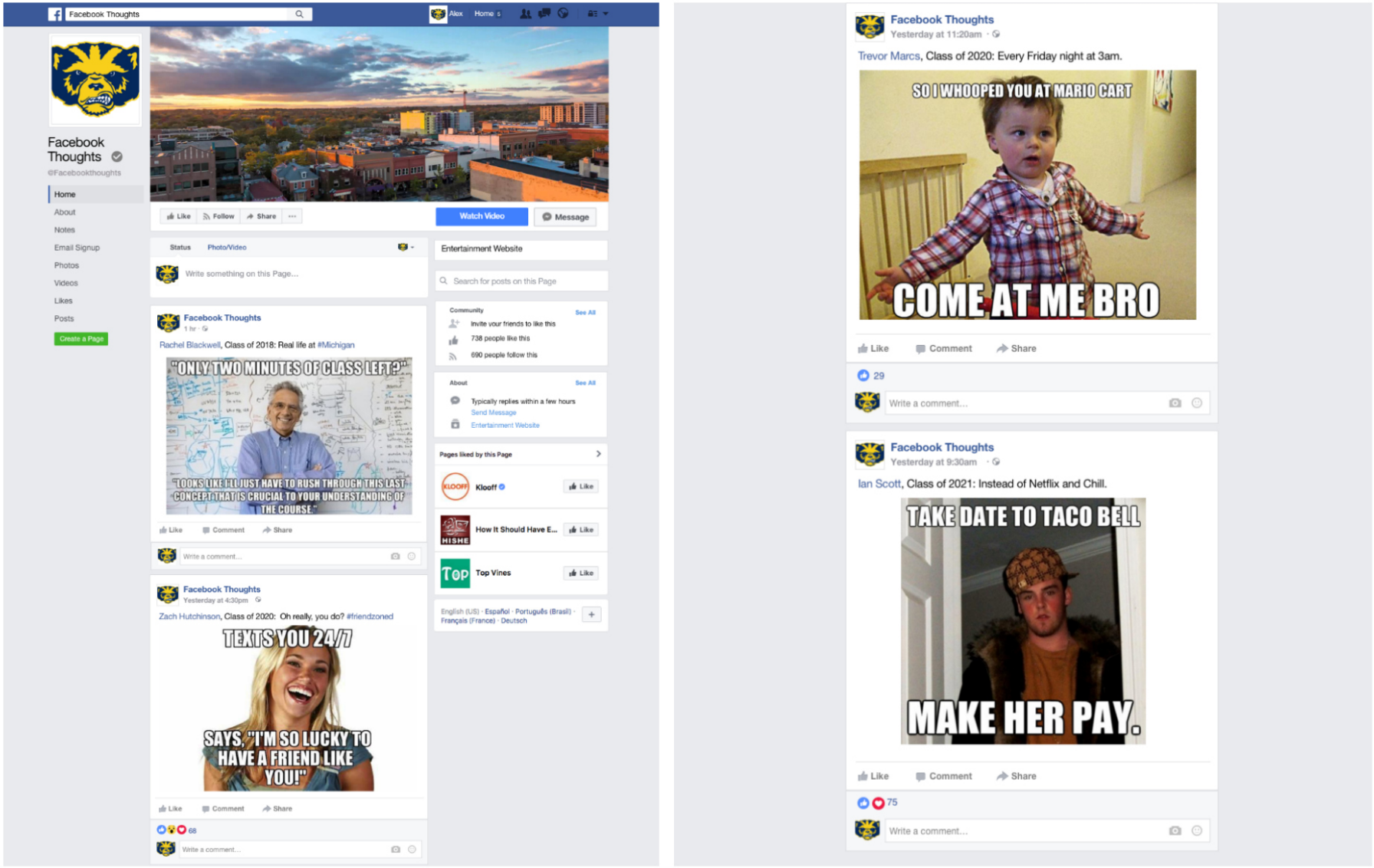
7. Winning is not my first priority. (reverse)
8. I enjoy taking risks.
9. I am disgusted by any kind of violence. (reverse)
10. I ask for help when I need it. (reverse)
11. My work is the most important part of my life.
12. I would only have sex if I was in a committed relationship. (reverse)
13. I bring up my feelings when talking to others. (reverse)
14. I would be furious if someone thought I was gay.
15. I don't mind losing. (reverse)
16. I take risks.
17. It would not bother me at all if someone thought I was gay. (reverse)
18. I never share my feelings.
19. Sometimes violent action is necessary.
20. In general, I control the women in my life.
21. I would feel good if I had many sexual partners.
22. It is important for me to win.
23. I don't like giving all my attention to work. (reverse)
24. It would be awful if people thought I was gay.
25. I like to talk about my feelings. (reverse)
26. I never ask for help.
27. More often than not, losing does not bother me. (reverse)
28. I frequently put myself in risky situations.
29. Women should be subservient to men.
30. I am willing to get into a physical fight if necessary.
31. I feel good when work is my first priority.
32. I tend to keep my feelings to myself.
33. Winning is not important to me. (reverse)
34. Violence is almost never justified. (reverse)
35. I am happiest when I'm risking danger.
36. It would be enjoyable to date more than one person at a time.
37. I would feel uncomfortable if someone thought I was gay.
38. I am not ashamed to ask for help. (reverse)
39. Work comes first.
40. I tend to share my feelings. (reverse)
41. No matter what the situation I would never act violently. (reverse)
42. Things tend to be better when men are in charge.
43. It bothers me when I have to ask for help.
44. I love it when men are in charge of women.
45. I hate it when people ask me to talk about my feelings.
46. I try to avoid being perceived as gay.

Appendix E

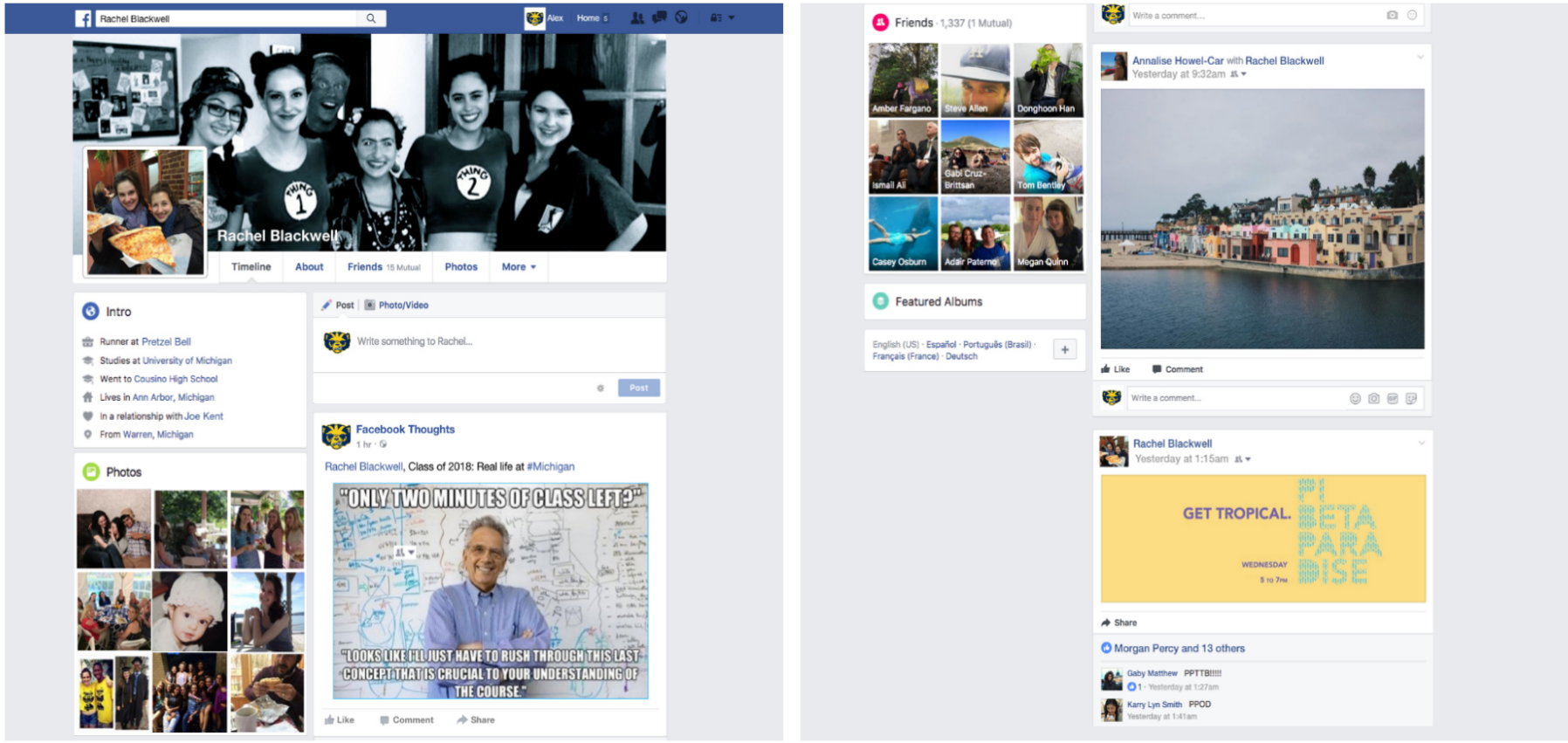
Figure 14. Facebook Thoughts pages (anonymous and identified conditions)



Anonymous Condition



Identified Condition



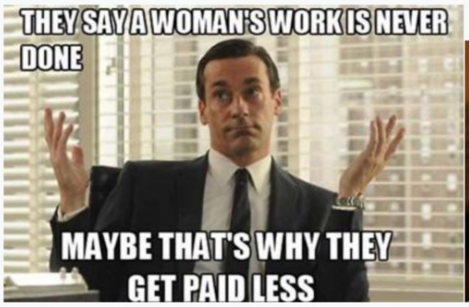
Identified Condition, “Tagged" participant

Appendix F

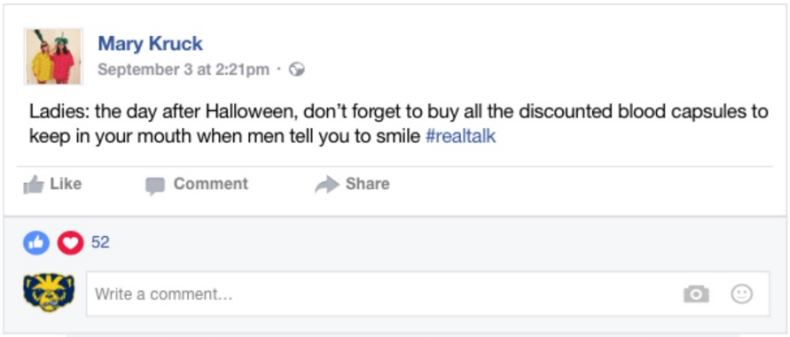
Figure 15. Example meme pairs for the meme selection task

Gender Equality Content

Please select ONE meme that you would post in response to the above status update.

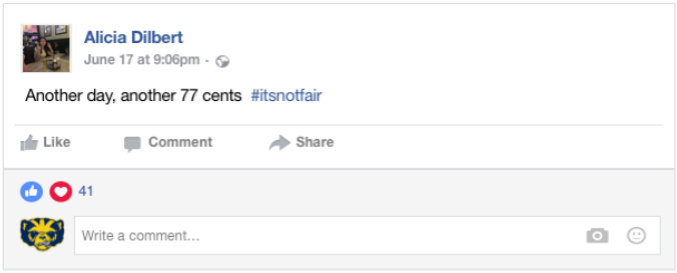


Please select ONE meme that you would post in response to the above status update.



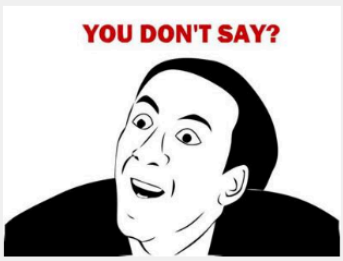
Please select ONE meme that you would post in response to the above status update.

Please select ONE meme that you would post in response to the above status update.



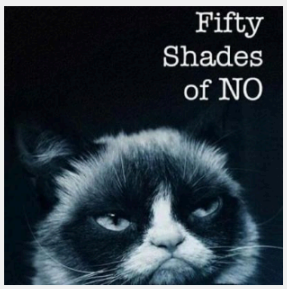
Please select ONE meme that you would post in response to the above status update.

Please select ONE meme that you would post in response to the above status update.



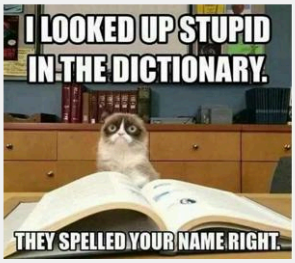
Please select ONE meme that you would post in response to the above status update.

Please select ONE meme that you would post in response to the above status update.

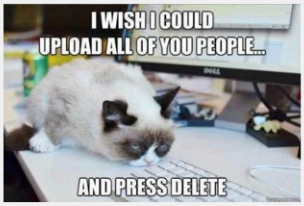


Neutral Content

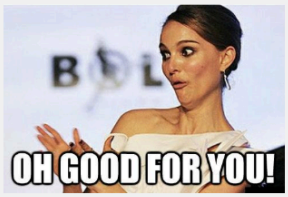
Please select ONE meme that you would post in response to the above status update.

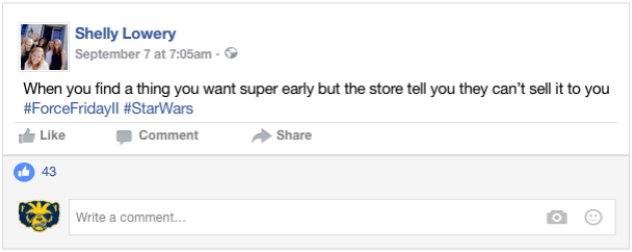


Please select ONE meme that you would post in response to the above status update.

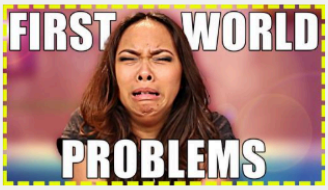


Please select ONE meme that you would post in response to the above status update.

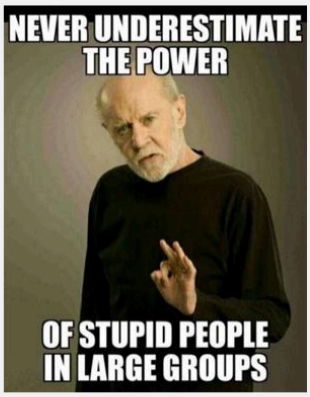


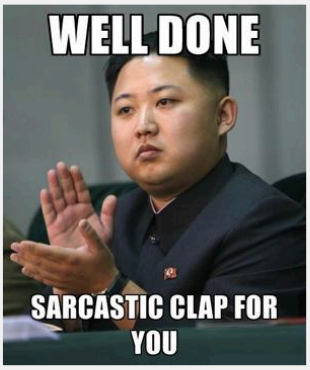


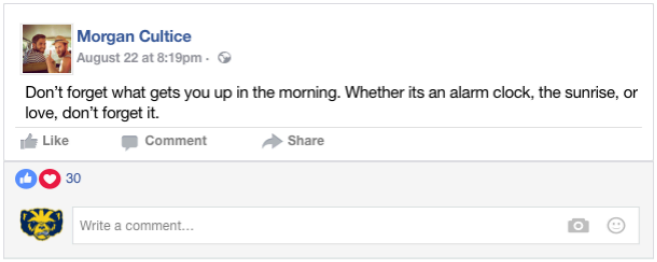
Please select ONE meme that you would post in response to the above status update.



Please select ONE meme that you would post in response to the above status update.

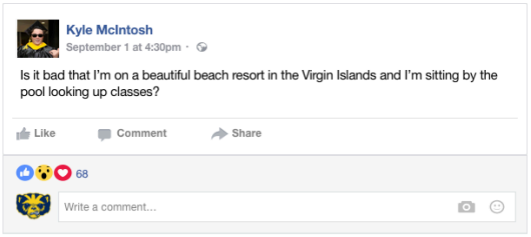




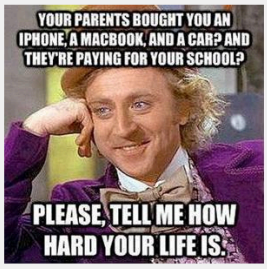


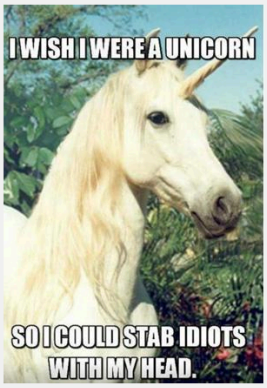
Please select ONE meme that you would post in response to the above status update.

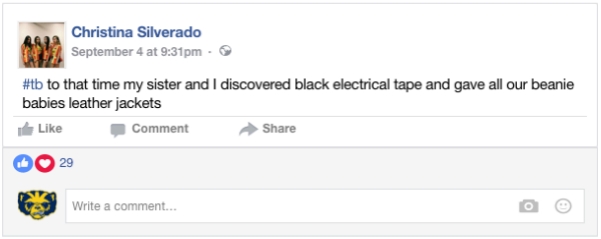
Please select ONE meme that you would post in response to the above status update.



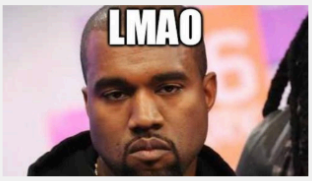
Please select ONE meme that you would post in response to the above status update.

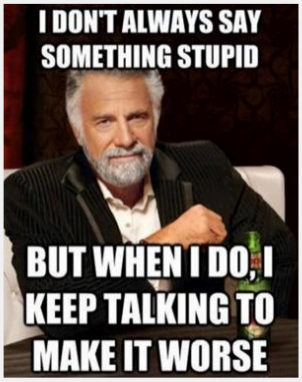






Please select ONE meme that you would post in response to the above status update.





Please select ONE meme that you would post in response to the above status update.

Appendix G

Study 3 Measures

Fearing Backlash

1=disagree strongly 10=agree strongly

Now imagine that your test result was public knowledge. Answer the following questions with this in mind.

1. Do you think you would feel proud?
2. Do you think you would be embarrassed?
3. Would your friends be likely to (negatively) tease you?
4. Would your family members be proud of you?
5. Would men give you a hard time (e.g., call you names)?
6. Would women give you a hard time (e.g., call you names)?
7. Would you be concerned that you might be disliked?
8. Would you worry about being labeled negatively?
9. Would you claim that the test was invalid?
10. Would you be concerned that people might think you're odd?

Facebook Intensity Scale

1=disagree strongly 2 =disagree somewhat 3 = disagree slightly 4 =agree slightly 5 =agree somewhat 6 =agree strongly

1. Facebook is part of my everyday activity.
2. I am proud to tell people that I am on Facebook.
3. Facebook has become part of my daily routine.
4. I feel out of touch when I haven't logged onto Facebook in awhile.
5. I feel I am part of the Facebook community.
6. I would be sorry if Facebook shut down.

Facebook Media Use

0-100 minutes, sliding scale

1. In the past week, on average, approximately how many MINUTES PER DAY have you spent actively using Twitter?

2. In the past week, on average, approximately how many MINUTES PER DAY have you spent:

- Posting Original Content in Facebook
- Reading Facebook Posts
- Responding to Facebook Posts
- Sharing media content (e.g., videos, memes, news stories) in Facebook